GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 25, 2005, 22:08:33 ; Search time 5981 Seconds
(without alignments)
11341.001 Million cell updates/sec

Title: US-10-070-255-4
Perfect score: 1782
Sequence: 1 atgagcggatcacagaacaa.....taaattgctttcattccttg 1782

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	821	46.1	1981	3	BC024293	BC024293 Homo sapi
2	653.8	36.7	998	5	BQ064251	BQ064251 AGENCOURT
3	651.6	36.6	990	5	BQ060776	BQ060776 AGENCOURT
c 4	650	36.5	691	4	BM668241	BM668241 UI-E-CK1-
5	648.2	36.4	970	5	BQ065026	BQ065026 AGENCOURT
6	639.8	35.9	730	1	AU122438	AU122438 AU122438
7	634.2	35.6	829	4	BG757448	BG757448 602711092
8	627.8	35.2	835	4	BG758677	BG758677 602712923
9	627.8	35.2	1022	5	BQ057616	BQ057616 AGENCOURT
10	624.6	35.1	843	4	BG758425	BG758425 602712623
11	624.4	35.0	892	5	BQ057008	BQ057008 AGENCOURT
12	598	33.6	609	4	BM692924	BM692924 UI-E-CK1-
13	574.8	32.3	813	4	BG684262	BG684262 602635896
14	535	30.0	819	4	BG759681	BG759681 602713374
c 15	534.6	30.0	626	5	BQ184171	BQ184171 UI-E-EJ1-
16	521.6	29.3	620	5	BQ185857	BQ185857 UI-E-EJ1-
17	513	28.8	769	4	BG684980	BG684980 602636729
18	505.6	28.4	782	4	BG759163	BG759163 602710820
19	500.6	28.1	914	4	BG340557	BG340557 602462123
20	494.6	27.8	951	4	BG685372	BG685372 602637216
21	491.8	27.6	1064	5	BQ065934	BQ065934 AGENCOURT
22	490.4	27.5	744	9	AY399018	AY399018 Homo sapi
23	478.8	26.9	737	4	BF975008	BF975008 602244582
24	478.2	26.8	1075	5	BQ058038	BQ058038 AGENCOURT
25	471	26.4	1051	5	BQ061744	BQ061744 AGENCOURT
c 26	462.6	26.0	855	5	BX760749	BX760749 BX760749
27	461.2	25.9	576	2	AW974284	AW974284 EST386387
28	455.4	25.6	822	4	BM478075	BM478075 AGENCOURT
29	453.2	25.4	965	4	BG757633	BG757633 602711219
30	450.4	25.3	504	4	BM697878	BM697878 UI-E-DX0-
c 31	439	24.6	439	1	AI369689	AI369689 qy71g08.x
32	428.6	24.1	501	1	AL135642	AL135642 DKFZp762K
33	423.8	23.8	682	9	AY399019	AY399019 Pan trogl
34	406.4	22.8	419	1	AA593354	AA593354 nn06h08.s
35	395	22.2	744	9	AY399020	AY399020 Mus muscu
36	384.8	21.6	430	2	AW403206	AW403206 UI-HF-BK0
37	382.4	21.5	850	5	BQ642454	BQ642454 AGENCOURT
38	382	21.4	506	5	BX283539	BX283539 BX283539
39	373.4	21.0	522	2	AW044949	AW044949 um15h06.y
40	370.8	20.8	520	2	AW503340	AW503340 UI-HF-BN0
41	355	19.9	478	5	BX283759	BX283759 BX283759
42	354.6	19.9	847	7	CR579211	CR579211 CR579211
43	342.4	19.2	646	1	AL965834	AL965834 AL965834
44	329.6	18.5	498	4	BG759800	BG759800 602711706
45	326.4	18.3	658	2	BB623950	BB623950 BB623950

ALIGNMENTS

RESULT 1
BC024293
LOCUS BC024293 1981 bp mRNA linear HTC 04-MAR-2003
DEFINITION Homo sapiens, clone IMAGE:5088007, mRNA.
ACCESSION BC024293
VERSION BC024293.1 GI:22028130
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1981)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (26-FEB-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 42 Row: m Column: 22
This clone has the following problem: retained intron.
FEATURES Location/Qualifiers
source 1. .1981
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5088007"
/tissue_type="Lymph, lymphoma"

```
                     /clone_lib="NIH_MGC_99"
                     /lab_host="DH10B"
                     /note="Vector: pOTB7"
ORIGIN

  Query Match             46.1%;  Score 821;  DB 3;  Length 1981;
  Best Local Similarity   97.3%;  Pred. No. 4.9e-210;
  Matches  857;  Conservative    0;  Mismatches   20;  Indels    4;  Gaps    2;

Qy        245 CAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGC-AGCGC 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        362 CAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAAGCGC 421

Qy        304 TTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGCTGT 363
              ||||||||||||||||||||||||||||||||||||||||| ||| ||||||||||||||
Db        422 TTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGCCGCACCTGGCTGCGCTGT 481

Qy        364 GCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGCAGG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 GCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGCAGG 541

Qy        424 CTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTT 483
              || || |||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        542 CTCAGTACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAGAGGTCCAGTATGCTT 601

Qy        484 CCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAGGAT 543
              |||||||||||||||||| |||||||||||||||||||||||||| ||||||||||||
Db        602 CCTACCATGGCAGCAGGTCCGAACTCCATACTCTTTGCGATTAACATTGACAACAAGGAT 661

Qy        544 TTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAG 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 TTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAG 721

Qy        604 AACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATC 663
              || ||||  |||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        722 AATGTGA---CCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCGTGTTCAGGGAGATC 778

Qy        664 ACAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCT 723
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        779 ACAGCCTCCTCTGCCATCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCC 838

Qy        724 GTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAG 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        839 GTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAG 898

Qy        784 AAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTG 843
              |||||||||||| |||||||||||||||||||||||||||||||||||||||||| ||
Db        899 CAAGTGACCAACATTATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACATG 958

Qy        844 TTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTC 903
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        959 TTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCACTCCTCTGTC 1018

Qy        904 AATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGCAGC 963
              |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db       1019 AATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCAAACTCCAATGGAAGTCAGAGCAGC 1078

Qy        964 AACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTTGAT 1023
              ||||| |||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db       1079 AACTCGTGGAAAATTGATTCCCTGTCTTTGAACAGGGAGTTTGGGTACCAGAAGCTTGAT 1138

Qy       1024 GTGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCA 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 GTGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCA 1198

Qy       1084 TCAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAA 1124
              |||||| |||||||||||||||||| ||||||||||||||
Db       1199 TCAGGACGGAAGCACAGGGGCCACTCAGAGTCGCCCGAGAA 1239


RESULT 2
BQ064251
LOCUS       BQ064251                998 bp    mRNA    linear   EST 02-APR-2002
```

```
DEFINITION  AGENCOURT_6854182 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5925960
            5', mRNA sequence.
ACCESSION   BQ064251
VERSION     BQ064251.1  GI:19892719
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 998)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2098  row: i  column: 01
            High quality sequence stop: 692.
FEATURES             Location/Qualifiers
     source          1. .998
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5925960"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_99"
                     /note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
                     of Gerald M. Rubin (University of California, Berkeley)
                     using ZAP-cDNA synthesis kit (Stratagene) and Superscript
                     II RT (Life Technologies). Note: this is a NIH_MGC
                     Library."
ORIGIN

  Query Match             36.7%;  Score 653.8;  DB 5;  Length 998;
  Best Local Similarity   97.0%;  Pred. No. 5.9e-165;
  Matches  677;  Conservative    0;  Mismatches   18;  Indels    3;  Gaps    1;

Qy        427 AGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTTCCT 486
              || |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         92 AGTACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAGAGGTCCAGTATGCTTCCT 151

Qy        487 ACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAGGATTTG 546
              |||||||||||||||| |||||||||||||||||||||||||| ||||||||||||||||
Db        152 ACCATGGCAGCAGGTCCGAACTCCATACTCTTTGCGATTAACATTGACAACAAGGATTTG 211

Qy        547 AACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAAC 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 AACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAAT 271

Qy        607 GTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCACA 666
              ||||   |||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        272 GTGA---CCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCGTGTTCAGGGAGATCACA 328

Qy        667 GCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCTGTC 726
              |||||||||||| |||||||||||||||||||||||||||||||||||||||| |||
Db        329 GCCTCCTCTGCCATCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCCGTC 388

Qy        727 GTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGAAA 786
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        389 GTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGCAA 448

Qy        787 GTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGTTT 846
```

```
                |||||||||||| |||||||||||||||||||||||||||||||||||||||||| |||||
Db        449 GTGACCAACATTATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACATGTTT 508

Qy        847 AAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCAAT 906
                |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        509 AAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCACTCCTCTGTCAAT 568

Qy        907 ATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGCAGCAAC 966
                |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        569 ATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCAAACTCCAATGGAAGTCAGAGCAGCAAC 628

Qy        967 TCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTTGATGTG 1026
                || |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        629 TCGTGGAAAATTGATTCCCTGTCTTTGAACAGGGAGTTTGGGTACCAGAAGCTTGATGTG 688

Qy       1027 AAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCATCA 1086
                |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        689 AAAAGCATCGATGATGAAGATGTGGATGANAACGAAGATGACGTGTATGGAAACTCATCA 748

Qy       1087 GGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAA 1124
                ||| |||||||||||||||||| ||||||||||||||
Db        749 GGACGGAAGCACAGGGGCCACTCAGAGTCGCCCGAGAA 786
```

RESULT 3
BQ060776

```
LOCUS       BQ060776                990 bp    mRNA    linear   EST 02-APR-2002
DEFINITION  AGENCOURT_7049038 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5816452
            5', mRNA sequence.
ACCESSION   BQ060776
VERSION     BQ060776.1  GI:19884191
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 990)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2073  row: f  column: 05
            High quality sequence stop: 686.
FEATURES             Location/Qualifiers
     source          1. .990
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5816452"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_99"
                     /note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
                     of Gerald M. Rubin (University of California, Berkeley)
                     using ZAP-cDNA synthesis kit (Stratagene) and Superscript
                     II RT (Life Technologies). Note: this is a NIH_MGC
                     Library."
ORIGIN
```

```
  Query Match           36.6%;  Score 651.6;  DB 5;  Length 990;
  Best Local Similarity  96.8%;  Pred. No. 2.3e-164;
  Matches  676;  Conservative    0;  Mismatches   19;  Indels    3;  Gaps    1;

Qy        427 AGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTTCCT 486
              || |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         98 AGTACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAGAGGTCCAGTATGCTTCCT 157

Qy        487 ACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAGGATTTG 546
              |||||||||||||||| |||||||||||||||||||||||||| |||||||||||||||
Db        158 ACCATGGCAGCAGGTCCGAACTCCATACTCTTTGCGATTAACATTGACAACAAGGATTTG 217

Qy        547 AACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAAC 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 AACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAAT 277

Qy        607 GTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCACA 666
              ||||   |||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        278 GTGA---CCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCGTGTTCAGGGAGATCACA 334

Qy        667 GCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCTGTC 726
              |||||||||||| |||||||||||||||||||||||||||||||||||||||||| |||
Db        335 GCCTCCTCTGCCATCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCCGTC 394

Qy        727 GTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGAAA 786
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        395 GTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGCAA 454

Qy        787 GTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGTTT 846
              |||||||||| |||||||||||||||||||||||||||||||||||||||||| |||||
Db        455 GTGACCAACATTATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACATGTTT 514

Qy        847 AAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCAAT 906
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        515 AAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCACTCCTCTGTCAAT 574

Qy        907 ATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGCAGCAAC 966
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        575 ATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCAAACTCCAATGGAAGTCAGAGCAGCAAC 634

Qy        967 TCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTTGATGTG 1026
              || |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        635 TCGTGGAAAATTGATTCCCTGTCTTTGAACAGGGAGTTTGGGTACCAGAAGCTTGATGTG 694

Qy       1027 AAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCATCA 1086
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        695 AAAAGCATCGATGATGAAGATCTGGATGAAAACGAAGATGACGTGTATGGAAACTCATCA 754

Qy       1087 GGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAA 1124
                || |||||||||||||||||| ||||||||||||||
Db        755 AGACGGAAGCACAGGGGCCACTCAGAGTCGCCCGAGAA 792
```

RESULT 4
BM668241/c

```
LOCUS       BM668241                691 bp    mRNA    linear   EST 27-FEB-2002
DEFINITION  UI-E-CK1-afk-h-18-0-UI.s2 UI-E-CK1 Homo sapiens cDNA clone
            UI-E-CK1-afk-h-18-0-UI 3', mRNA sequence.
ACCESSION   BM668241
VERSION     BM668241.1  GI:18976072
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 691)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
```

```
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Forward
            POLYA=Yes.
FEATURES              Location/Qualifiers
     source           1. .691
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="UI-E-CK1-afk-h-18-0-UI"
                      /tissue_type="Retina Foveal and Macular"
                      /dev_stage="adult"
                      /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                      /clone_lib="UI-E-CK1"
                      /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                      modified polylinker; Site_1: EcoR I; Site_2: Not I;
                      UI-E-CK1 is a normalized cDNA library containing the
                      following tissue(s): Retina Foveal and Macular. The
                      library was constructed according to Bonaldo, Lennon and
                      Soares, Genome Research, 6:791-806, 1996. First strand
                      cDNA synthesis was primed with an oligo-dT primer
                      containing a Not I site. Double stranded cDNA was ligated
                      to an EcoR I adaptor, digested with Not I, and cloned
                      directionally into pT7T3-Pac vector. The oligonucleotide
                      used to prime the synthesis of first-strand cDNA contains
                      a library tag sequence that is located between the Not I
                      site and the (dT)18 tail. The sequence tag for this
                      library is GTCC. This library was created for the program,
                      Gene Discovery in the Visual System, supported by National
                      Eye Institute (NEI).
                      TAG_TISSUE=Foveal and Macular Retina
                      TAG_LIB=UI-E-CK1
                      TAG_SEQ=GTCC"
ORIGIN

  Query Match           36.5%;  Score 650;  DB 4;  Length 691;
  Best Local Similarity   99.7%;  Pred. No. 5.6e-164;
  Matches  672;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    2;

Qy       1110 GGAGTCGCCCGAGAAGCCACTGGAAGGGAACACCTGCCTCTCCCAGATGCACAGCTGGGC 1169
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        691 GGAGTCGCCCGAGAAGCCACTGGAAGGGAACACCTGCCTCTCCCAGATGCACAGCTGGGC 632

Qy       1170 TCCGCTGAAGGTGCTGCACAATGACTCCGACATCCTCTTCCCTGTCAGTGGCGTGGGCTC 1229
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        631 TCCGCTGAAGGTGCTGCACAATGACTCCGACATCCTCTTCCCTGTCAGTGGCGTGGGCTC 572

Qy       1230 CTACAGCCCAGCAGATGCCCCCCTCGGAAGCCTGGAGAACGGGACAGGACCAGAGGACCA 1289
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        571 CTACAGCCCAGCAGATGCCCCCCTCGGAAGCCTGGAGAACGGGACAGGACCAGAGGACCA 512

Qy       1290 CGTTCTCCCGGATCCTGGACTTCGGTACAGTGTGGAAGCCAGCTCTCCAGGCCACGGAAG 1349
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 CGTTCTCCCGGATCCTGGACTTCGGTACAGTGTGGAAGCCAGCTCTCCAGGCCACGGAAG 452

Qy       1350 TCCTCTGAGCAGCCTGTTACCTTCTGCCTCAGTGCCAGAGTCCATGACAATTAGTGAACT 1409
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 TCCTCTGAGCAGCCTGTTACCTTCTGCCTCAGTGCCAGAGTCCATGACAATTAGTGAACT 392

Qy       1410 GCGCCAGGCCACTGTGGCCATGATGAACAGGAAGGATGAGCTGGAGGAGGAGAACAGATC 1469
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 GCGCCAGGCCACTGTGGCCATGATGAACAGGAAGGATGAGCTGGAGGAGGAGAACAGATC 332

Qy       1470 ACTGCGAAACCTGCTCGACGGTGAGATGGAGCACTCAGCCGCGCTCCGGCAAGAGGTGGA 1529
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 ACTGCGAAACCTGCTCGACGGTGAGATGGAGCACTCAGCCGCGCTCCGGCAAGAGGTGGA 272

Qy       1530 CACCTTGAAAAGGAAGGTGGCTGAACAGGAGGAGCGGCAGGGCATGAAGGTCCAGGCGCT 1589
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 CACCTTGAAAAGGAAGGTGGCTGAACAGGAGGAGCGGCAGGGCATGAAGGTCCAGGCGCT 212

Qy       1590 GGCCAGCTATCTTTGCTATTTTGTGAGGAGATTCTAACCCCACGTGAGAACCATGTGGTG 1649
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 GGCCAGCTATCTTTGCTATTTTGTGAGGAGATTCTAACCCCACGTGAGAACCATGTGGTG 152

Qy       1650 GAGAAATGGAGGGAGAGAGAAATCCAACAGTTCCTGATAGTCTCATTTGAGCTCCTGGAT 1709
             ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        151 GAGAAATGGAGGGAGAGAGAAAT-CAACAGTTCCTGATAGTCTCATTTGAGCTCCTGGAT 93

Qy       1710 CCAGTCTTTCCTGAAGCTGTGTTTCCTCTGGACTTTTCATGTATGTGAG-CCAATAAATT 1768
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db          92 CCAGTCTTTCCTGAAGCTGTGTTTCCTCTGGACTTTTCATGTATGTGAGCCCAATAAATT 33

Qy       1769 GCTTTCATTCCTTG 1782
             ||||||||||||||
Db          32 GCTTTCATTCCTTG 19


RESULT 5
BQ065026
LOCUS       BQ065026                970 bp    mRNA    linear   EST 02-APR-2002
DEFINITION  AGENCOURT_6856627 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5929505
            5', mRNA sequence.
ACCESSION   BQ065026
VERSION     BQ065026.1  GI:19894072
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 970)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2107  row: l  column: 18
            High quality sequence stop: 676.
FEATURES              Location/Qualifiers
     source           1. .970
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:5929505"
                      /tissue_type="lymphoma, cell line"
                      /lab_host="DH10B (phage-resistant)"
                      /clone_lib="NIH_MGC_99"
                      /note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
                      EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                      into EcoRI/XhoI sites using the following 5' adaptor:
                      GGCACGAG(G). Size-selected >500bp for average insert size
                      1.8kb. Library constructed by Ling Hong in the laboratory
                      of Gerald M. Rubin (University of California, Berkeley)
                      using ZAP-cDNA synthesis kit (Stratagene) and Superscript
                      II RT (Life Technologies). Note: this is a NIH_MGC
                      Library."
ORIGIN

  Query Match           36.4%;  Score 648.2;  DB 5;  Length 970;
  Best Local Similarity   97.0%;  Pred. No. 1.9e-163;
  Matches  672;  Conservative    0;  Mismatches   18;  Indels    3;  Gaps    1;
```

```
Qy        427 AGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTTCCT 486
              || |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         92 AGTACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAGAGGTCCAGTATGCTTCCT 151

Qy        487 ACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAGGATTTG 546
              |||||||||||||||| |||||||||||||||||||||||| ||||||||||||||||||
Db        152 ACCATGGCAGCAGGTCCGAACTCCATACTCTTTGCGATTAACATTGACAACAAGGATTTG 211

Qy        547 AACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAAC 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 AACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAAT 271

Qy        607 GTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCACA 666
              ||||   ||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        272 GTGA---CCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCGTGTTCAGGGAGATCACA 328

Qy        667 GCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCTGTC 726
              |||||||||||| |||||||||||||||||||||||||||||||||||||||||||| |||
Db        329 GCCTCCTCTGCCATCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCCGTC 388

Qy        727 GTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGAAA 786
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        389 GTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGCAA 448

Qy        787 GTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGTTT 846
              |||||||||||| |||||||||||||||||||||||||||||||||||||| |||||
Db        449 GTGACCAACATTATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACATGTTT 508

Qy        847 AAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCAAT 906
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        509 AAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCACTCCTCTGTCAAT 568

Qy        907 ATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGCAGCAAC 966
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        569 ATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCAAACTCCAATGGAAGTCAGAGCAGCAAC 628

Qy        967 TCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTTGATGTG 1026
              || |||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        629 TCGTGGAAAATTGATTCCCTGTCTTTGAACAGGGAGTTTGGGTACCAGAAGCTTGATGTG 688

Qy       1027 AAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCATCA 1086
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 AAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCATCA 748

Qy       1087 GGAAGGAAGCACAGGGGCCACTCGGAGTCGCCC 1119
              ||| |||||||||||||||||  ||||||||
Db        749 GGACGGAAGCACAGGGGCCACTCAAAGTCGCCC 781
```

RESULT 6
AU122438

```
LOCUS       AU122438                730 bp    mRNA    linear   EST 01-AUG-2002
DEFINITION  AU122438 MAMMA1 Homo sapiens cDNA clone MAMMA1002380 5', mRNA
            sequence.
ACCESSION   AU122438
VERSION     AU122438.1  GI:10937708
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 730)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .730
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MAMMA1002380"
                     /tissue_type="mammary gland"
                     /clone_lib="MAMMA1"
                     /note="Vector: pME18SFL3"
ORIGIN
```

Query Match 35.9%; Score 639.8; DB 1; Length 730;
Best Local Similarity 99.0%; Pred. No. 3.3e-161;
Matches 663; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

```
Qy          1 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 122

Qy         61 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 182

Qy        121 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 242

Qy        181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 302

Qy        241 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 362

Qy        301 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 422

Qy        361 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 482

Qy        421 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 542

Qy        481 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 602

Qy        541 GATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 600
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 GATTTGGACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 662

Qy        601 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 660
              || ||||||| || |||||||||||||||||||||||||||| || |||||||| ||||
Db        663 CANAACGTGA-CTNCTTGCTGAAGGAGTCCACGCAAGGAGTGANCA-CCTGTTCAAGGAG 720

Qy        661 ATCACAGCCT 670
              ||||||||||
Db        721 ATCACAGCCT 730
```

RESULT 7
BG757448

```
LOCUS       BG757448                829 bp    mRNA    linear   EST 15-MAY-2001
```

```
DEFINITION  602711092F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4851502 5',
            mRNA sequence.
ACCESSION   BG757448
VERSION     BG757448.1  GI:14068101
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 829)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1694  row: g  column: 23
            High quality sequence stop: 801.
FEATURES             Location/Qualifiers
     source          1. .829
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4851502"
                     /tissue_type="primary B-cells from tonsils (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_48"
                     /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by Ling
                     Hong in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             35.6%;  Score 634.2;  DB 4;  Length 829;
  Best Local Similarity   96.1%;  Pred. No. 1.1e-159;
  Matches  672;  Conservative    0;  Mismatches   23;  Indels    4;  Gaps    2;

Qy        427 AGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTTCCT 486
              || |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        133 AGTACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAGAGGTCCAGTATGCTTCCT 192

Qy        487 ACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAGGATTTG 546
              |||||||||||||||| |||||||||||||||||||||||| |||||||||||||||||
Db        193 ACCATGGCAGCAGGTCCGAACTCCATACTCTTTGCGATTAACATTGACAACAAGGATTTG 252

Qy        547 AACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAAC 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 AACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAAT 312

Qy        607 GTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCACA 666
              ||||   |||||||||||||||||||||||||||||||| ||||||| ||||||||||||
Db        313 GTGA---CCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCGTGTTCAGGCAGATCACA 369

Qy        667 GCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCTGTC 726
              |||||||||||| |||||||||||||||||||||||||||||||||||||||||| |||
Db        370 GCCTCCTCTGCCATCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCCGTC 429

Qy        727 GTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGAAA 786
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        430 GTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGCAA 489

Qy        787 GTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGTTT 846
              |||||||||| |||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        490 GTGACCAACATTATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACATGTTT 549

Qy        847 AAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCAAT 906
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        550 AAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCACTCCTCTGTCAAT 609

Qy        907 ATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGCAGCAAC 966
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        610 ATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCAAACTCCAATGGAAGTCAGAGCAGCAAC 669

Qy        967 TCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTTGATGTG 1026
              || |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        670 TCGTGGAAAATTGATTCCCTGTCTTTGAACAGGGAGTTTGGGTACCAGAAGCTTGATGTG 729

Qy       1027 AAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGG-AAACTCATC 1085
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        730 AAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAAACTCATC 789

Qy       1086 AGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAA 1124
              |||| ||||||||| | | ||||| ||||||  ||||||
Db        790 AGGACGGAAGCACATGTGGCACTCAGAGTCGGGCGAGAA 828


RESULT 8
BG758677
LOCUS       BG758677                 835 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602712923F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4853188 5',
            mRNA sequence.
ACCESSION   BG758677
VERSION     BG758677.1  GI:14069330
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 835)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1698  row: n  column: 05
            High quality sequence stop: 819.
FEATURES             Location/Qualifiers
     source          1. .835
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4853188"
                     /tissue_type="primary B-cells from tonsils (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_48"
                     /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by Ling
                     Hong in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             35.2%;  Score 627.8;  DB 4;  Length 835;
  Best Local Similarity   96.6%;  Pred. No. 5.9e-158;
```

```
  Matches  675;  Conservative    0;  Mismatches   17;  Indels    7;  Gaps    3;

Qy        427 AGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTTCCT 486
              || |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         31 AGTACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAGAGGTCCAGTATGCTTCCT 90

Qy        487 ACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAGGATTTG 546
              |||||||||||||||| ||||||||||||||||||||   || ||||||||||||||||
Db         91 ACCATGGCAGCAGGTCCGAACTCCATACTCTTTGCGATT---ATTGACAACAAGGATTTG 147

Qy        547 AACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAAC 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 AACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAAT 207

Qy        607 GTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCACA 666
              ||||   |||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        208 GTGA---CCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCGTGTTCAGGGAGATCACA 264

Qy        667 GCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCTGTC 726
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||| |||
Db        265 GCCTCCTCTGCCATCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCCGTC 324

Qy        727 GTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGAAA 786
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        325 GTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGCAA 384

Qy        787 GTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGTTT 846
              |||||||||||| |||||||||||||||||||||||||||||||||||||||||| |||||
Db        385 GTGACCAACATTATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACATGTTT 444

Qy        847 AAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCAAT 906
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        445 AAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCACTCCTCTGTCAAT 504

Qy        907 ATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGCAGCAAC 966
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        505 ATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCAAACTCCAATGGAAGTCAGAGCAGCAAC 564

Qy        967 TCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTTGATGTG 1026
              || |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        565 TCGTGGAAAATTGATTCCCTGTCTTTGAACAGGGAGTTTGGGTACCAGAAGCTTGATGTG 624

Qy       1027 -AAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCATC 1085
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        625 AAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCATC 684

Qy       1086 AGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAA 1124
              |||| ||||||||||||||||||| ||||||||||||||
Db        685 AGGACGGAAGCACAGGGGCCACTCAGAGTCGCCCGAGAA 723


RESULT 9
BQ057616
LOCUS       BQ057616               1022 bp    mRNA    linear   EST 29-MAR-2002
DEFINITION  AGENCOURT_7047416 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5813279
            5', mRNA sequence.
ACCESSION   BQ057616
VERSION     BQ057616.1  GI:19816956
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1022)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2065  row: a  column: 24
            High quality sequence stop: 641.
FEATURES             Location/Qualifiers
     source          1. .1022
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5813279"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_99"
                     /note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
                     of Gerald M. Rubin (University of California, Berkeley)
                     using ZAP-cDNA synthesis kit (Stratagene) and Superscript
                     II RT (Life Technologies). Note: this is a NIH_MGC
                     Library."
ORIGIN

  Query Match             35.2%;  Score 627.8;  DB 5;  Length 1022;
  Best Local Similarity   96.3%;  Pred. No. 6.3e-158;
  Matches  654;  Conservative    0;  Mismatches   22;  Indels    3;  Gaps    1;

Qy        427 AGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTTCCT 486
              || |||||||||||||||||||||||||||||||||||||||||| || ||||||||||||||||
Db         70 AGTACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAGAGTTCCAGTATGCTTCCT 129

Qy        487 ACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAGGATTTG 546
              |||||||||||||||| |||||||||||||||||||||||||||| ||||||||||||||||
Db        130 ACCATGGCAGCAGGTCCGAACTCCATACTCTTTGCGATTAACATTGACAACAAGGATTTG 189

Qy        547 AACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAAC 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 AACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAAT 249

Qy        607 GTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCACA 666
              ||||   |||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        250 GTGA---CCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCGTGTTCAGGGAGATCACA 306

Qy        667 GCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCTGTC 726
              |||||||||||| |||||||||||||||||||||||||||||||||||||||||| |||
Db        307 GCCTCCTCTGCCATCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCCGTC 366

Qy        727 GTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGAAA 786
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        367 GTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGCAA 426

Qy        787 GTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGTTT 846
              |||||||||||| |||||||||||||||||||||||||||||||||||||||||| |||||
Db        427 GTGACCAACATTATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACATGTTT 486

Qy        847 AAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCAAT 906
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        487 AAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCACTCCTCTGTCAAT 546

Qy        907 ATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGCAGCAAC 966
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        547 ATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCAAACTCCAATGGAAGTCAGAGCAGCAAC 606

Qy        967 TCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTTGATGTG 1026
              || |||||||||||||||||||||||||| ||||||||||| ||||||||||||||||||||
Db        607 TCGTGGAAAATTGATTCCCTGTCTTTGAACAGGGAGTTTGGGGACCAGAAGCTTGATGTG 666

Qy       1027 AAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCATCA 1086
              |||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||
Db        667 AAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACTGGTATGGAAACTCATCA 726
```

```
Qy       1087 GGAAGGAAGCACAGGGGCC 1105
              ||| |||| |||||||| |
Db        727 GGACGGAAACACAGGGGGC 745


RESULT 10
BG758425
LOCUS       BG758425                843 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602712623F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4853164 5',
            mRNA sequence.
ACCESSION   BG758425
VERSION     BG758425.1  GI:14069078
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 843)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1698  row: m  column: 05
            High quality sequence stop: 835.
FEATURES             Location/Qualifiers
     source          1. .843
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4853164"
                     /tissue_type="primary B-cells from tonsils (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_48"
                     /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by Ling
                     Hong in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             35.1%;  Score 624.6;  DB 4;  Length 843;
  Best Local Similarity   96.5%;  Pred. No. 4.4e-157;
  Matches  661;  Conservative    0;  Mismatches   19;  Indels    5;  Gaps    2;

Qy        427 AGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTTCCT 486
              || |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        142 AGTACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAGAGGTCCAGTATGCTTCCT 201

Qy        487 ACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAGGATTTG 546
              ||||||||||||||| |||||||||||||||||||||||| ||||||||||||||||||
Db        202 ACCATGGCAGCAGGTCCGAACTCCATACTCTTTGCGATTAACATTGACAACAAGGATTTG 261

Qy        547 AACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAAC 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 AACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAAT 321

Qy        607 GTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCACA 666
              ||||   |||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        322 GTGA---CCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCGTGTTCAGGGAGATCACA 378

Qy        667 GCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCTGTC 726
              |||||||||||| |||||||||||||||||||||||||||||||||||||||||| |||
Db        379 GCCTCCTCTGCCATCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCCGTC 438

Qy        727 GTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGAAA 786
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        439 GTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGCAA 498

Qy        787 GTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGTTT 846
              |||||||||||| |||||||||||||||||||||||||||||||||||||||||| |||||
Db        499 GTGACCAACATTATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACATGTTT 558

Qy        847 AAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCAAT 906
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        559 AAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCACTCCTCTGTCAAT 618

Qy        907 ATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGCAGCAAC 966
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db        619 ATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCAAACTCCAATGGAAGTCAGAGCAGCAAC 678

Qy        967 TCATGG--AAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTTGATG 1024
              || |||  ||| | ||||||||||||||||| ||||||||||||||||||||||||||||
Db        679 TCGTGGGAAAATTGGATTCCCTGTCTTTGAACAGGGAGTTTGGGTACCAGAAGCTTGATG 738

Qy       1025 TGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCAT 1084
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        739 TGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGACACTCAT 798

Qy       1085 CAGGAAGGAAGCACAGGGGCCACTC 1109
              ||||| ||||||||||||||||||||
Db        799 CAGGACGGAAGCACAGGGGCCACTC 823


RESULT 11
BQ057008
LOCUS       BQ057008                892 bp    mRNA    linear   EST 29-MAR-2002
DEFINITION  AGENCOURT_6769190 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5812156
            5', mRNA sequence.
ACCESSION   BQ057008
VERSION     BQ057008.1  GI:19816348
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 892)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2062  row: c  column: 05
            High quality sequence stop: 608.
FEATURES             Location/Qualifiers
     source          1. .892
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5812156"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_99"
                     /note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
```

of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 35.0%; Score 624.4; DB 5; Length 892;
Best Local Similarity 92.4%; Pred. No. 5e-157;
Matches 680; Conservative 0; Mismatches 51; Indels 5; Gaps 2;

```
Qy        427 AGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTTCCT 486
              || |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         98 AGTACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAGAGGTCCAGTATGCTTCCT 157

Qy        487 ACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAGGATTTG 546
              |||||||||||||||| |||||||||||||||||||||||||||| ||||||||||||||
Db        158 ACCATGGCAGCAGGTCCGAACTCCATACTCTTTGCGATTAACATTGACAACAAGGATTTG 217

Qy        547 AACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAAC 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 AACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAAT 277

Qy        607 GTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCACA 666
              ||||   |||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        278 GTGA---CCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCGTGTTCAGGGAGATCACA 334

Qy        667 GCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCTGTC 726
              |||||||||||| |||||||||||||||||||||||||||||||||||||||||| |||
Db        335 GCCTCCTCTGCCATCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCCGTC 394

Qy        727 GTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGAAA 786
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        395 GTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGCAA 454

Qy        787 GTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGTTT 846
              |||||||||||| |||||||||||||||||||||||||||||||||||||||||| |||||
Db        455 GTGACCAACATTATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACATGTTT 514

Qy        847 AAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCAAT 906
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        515 AAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCACTCCTCTGTCAAT 574

Qy        907 ATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGCAGCAAC 966
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        575 ATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCAAACTCCAATGGAAGTCAGAGCAGCAAC 634

Qy        967 TCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTTGATGTG 1026
              || |||||||||||||||||||||||||| |||||||||||||||||| |||||||||
Db        635 TCGTGGAAAATTGATTCCCTGTCTTTGAACAGGGAGTTTGGGTACCAGAACCTTGATGTG 694

Qy       1027 AAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCATCA 1086
              |||||||| |||||||||| |||||||||| || ||||||||||||||||||||||||||
Db        695 AAAAGCATCCATGATGAAGATCTGGATGAAAACCAAAATGACGTGTATGGAAACTCATCA 754

Qy       1087 GGAA--GGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAAGCCACTGGAAGGGAACACCT 1144
              || |  | |||| |||||||||||| |||   |||   |     |  | |     ||  ||
Db        755 GGGACGGAAAGCCCAGGGGCCACTCCGAGATCCCCCGCAGAAAAAACGGAACACACTTCT 814

Qy       1145 GCCTCTCCCAGATGCA 1160
              | | | ||  | || |
Db        815 GTCACCCCAGGCTGGA 830
```

RESULT 12
BM692924
LOCUS BM692924 609 bp mRNA linear EST 28-FEB-2002
DEFINITION UI-E-CK1-afk-h-18-0-UI.r1 UI-E-CK1 Homo sapiens cDNA clone UI-E-CK1-afk-h-18-0-UI 5', mRNA sequence.
ACCESSION BM692924
VERSION BM692924.1 GI:19006182
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 609)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
Seq primer: M13 Reverse.
FEATURES Location/Qualifiers
source 1. .609
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-CK1-afk-h-18-0-UI"
/tissue_type="Retina Foveal and Macular"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-CK1"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-E-CK1 is a normalized cDNA library containing the following tissue(s): Retina Foveal and Macular. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GTCC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

ORIGIN

Query Match 33.6%; Score 598; DB 4; Length 609;
Best Local Similarity 99.8%; Pred. No. 6.1e-150;
Matches 609; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy       1100 GGGGCCACTCGGAGTCGCCCGAGAAGCCACTGGAAGGGAACACCTGCCTCTCCCAGATGC 1159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGGCCACTCGGAGTCGCCCGAGAAGCCACTGGAAGGGAACACCTGCCTCTCCCAGATGC 60

Qy       1160 ACAGCTGGGCTCCGCTGAAGGTGCTGCACAATGACTCCGACATCCTCTTCCCTGTCAGTG 1219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACAGCTGGGCTCCGCTGAAGGTGCTGCACAATGACTCCGACATCCTCTTCCCTGTCAGTG 120

Qy       1220 GCGTGGGCTCCTACAGCCCAGCAGATGCCCCCCCTCGGAAGCCTGGAGAACGGGACAGGAC 1279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGTGGGCTCCTACAGCCCAGCAGATGCCCCCCCTCGGAAGCCTGGAGAACGGGACAGGAC 180

Qy       1280 CAGAGGACCACGTTCTCCCGGATCCTGGACTTCGGTACAGTGTGGAAGCCAGCTCTCCAG 1339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGAGGACCACGTTCTCCCGGATCCTGGACTTCGGTACAGTGTGGAAGCCAGCTCTCCAG 240

Qy       1340 GCCACGGAAGTCCTCTGAGCAGCCTGTTACCTTCTGCCTCAGTGCCAGAGTCCATGACAA 1399
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        241 GCCACGGAAGTCCTCTGAGCAGCCTGTTACCTTCTGCCTCAGTGCCAGAGTCCATGACAA 300

Qy       1400 TTAGTGAACTGCGCCAGGCCACTGTGGCCATGATGAACAGGAAGGATGAGCTGGAGGAGG 1459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTAGTGAACTGCGCCAGGCCACTGTGGCCATGATGAACAGGAAGGATGAGCTGGAGGAGG 360

Qy       1460 AGAACAGATCACTGCGAAACCTGCTCGACGGTGAGATGGAGCACTCAGCCGCGCTCCGGC 1519
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGAACAGATCACTGCGAAACCTGCTCGACGGTGAGATGGAGCACTCAGCCGCGCTCCGGC 420

Qy       1520 AAGAGGTGGACACCTTGAAAAGGAAGGTGGCTGAACAGGAGGAGCGGCAGGGCATGAAGG 1579
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGAGGTGGACACCTTGAAAAGGAAGGTGGCTGAACAGGAGGAGCGGCAGGGCATGAAGG 480

Qy       1580 TCCAGGCGCTGGCCAGCTATCTTTGCTATTTTGTGAGGAGATTCTAACCCCACGTGAGAA 1639
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCCAGGCGCTGGCCAGCTATCTTTGCTATTTTGTGAGGAGATTCTAACCCCACGTGAGAA 540

Qy       1640 CCATGTGGTGGAGAAATGGAGGGAGAGAGAAATCCAACAGTTCCTGATAGTCTCATTTGA 1699
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        541 CCATGTGGTGGAGAAATGGAGGGAGAGAGAAAT-CAACAGTTCCTGATAGTCTCATTTGA 599

Qy       1700 GCTCCTGGAT 1709
              ||||||||||
Db        600 GCTCCTGGAT 609


RESULT 13
BG684262
LOCUS       BG684262                 813 bp    mRNA    linear   EST 01-MAY-2001
DEFINITION  602635896F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4763783 5',
            mRNA sequence.
ACCESSION   BG684262
VERSION     BG684262.1  GI:13915659
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 813)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1619  row: p  column: 24
            High quality sequence stop: 795.
FEATURES             Location/Qualifiers
     source          1. .813
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4763783"
                     /tissue_type="primary B-cells from tonsils (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_48"
                     /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by Ling
                     Hong in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN
```

```
  Query Match           32.3%;  Score 574.8;  DB 4;  Length 813;
  Best Local Similarity  96.3%;  Pred. No. 1.2e-143;
  Matches  653;  Conservative    0;  Mismatches   17;  Indels    8;  Gaps    6;

Qy        427 AGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTTCCT 486
              || |||||||||||||||||||| |||||||||||||||| ||||||||||||||||||||
Db        142 AGTACTTTTTATGAAGACTGGTC-TTTGTGATGGATGAAGAGAGGTCCAGTATGCTTCCT 200

Qy        487 ACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAGGATTTG 546
              |||||||||||||||| |||||||||||||||||||||||| ||||||||||||||||||
Db        201 ACCATGGCAGCAGGTCCGAACTCCATACTCTTTGCGATTAACATTGACAACAAGGATTTG 260

Qy        547 AACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAAC 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 AACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAAT 320

Qy        607 GTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCACA 666
              ||||   |||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        321 GTGA---CCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCGTGTTCAGGGAGATCACA 377

Qy        667 GCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCCGTC 726
              |||||||||||| |||||||||||||||||||||||||||||||||||||||| |||||
Db        378 GCCTCCTCTGCCATCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCCGTC 437

Qy        727 GTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGAAA 786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        438 GTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGCAA 497

Qy        787 GTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGTTT 846
              |||||||||||| |||||||||||||||||||||||||||||||||||||||| |||||
Db        498 GTGACCAACATTATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACATGTTT 557

Qy        847 AAAAAGA-CACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCAA 905
              ||||||| |||||||||||||||||||||||||||||||||||||  ||||||||||||
Db        558 AAAAAGACCACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACC-ATCCTCTGTCAA 616

Qy        906 TATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGCAGCAA 965
              ||||||||||||||||| |||||||||||||| |||||||||||||||||||||||||||
Db        617 TATCATGTCCGCCTTTGACAGCCCCTTCGGGCCAAACTCCAATGGAAGTCAGAGCAGCAA 676

Qy        966 CTCATGGAAAATT-GATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTTGATG 1024
              ||| ||||||||| |||||||||||||||| ||||| |||||||||||||||||||||||
Db        677 CTCGTGGAAAATTGGATTCCCTGTCTTTGAACAGGGAG-TTGGGTACCAGAAGCTTGATG 735

Qy       1025 TGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCAT 1084
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 TGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCAT 795

Qy       1085 CAGGAAGGAAGCACAGGG 1102
              ||||| ||||||||||||
Db        796 CAGGACGGAAGCACAGGG 813


RESULT 14
BG759681
LOCUS       BG759681                 819 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602713374F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4853346 5',
            mRNA sequence.
ACCESSION   BG759681
VERSION     BG759681.1  GI:14070334
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 819)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
```

```
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLCM1699  row: d  column: 19
           High quality sequence stop: 791.
FEATURES             Location/Qualifiers
     source          1. .819
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4853346"
                     /tissue_type="primary B-cells from tonsils (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_48"
                     /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by Ling
                     Hong in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             30.0%;  Score 535;  DB 4;  Length 819;
  Best Local Similarity   95.4%;  Pred. No. 7e-133;
  Matches  617;  Conservative    0;  Mismatches   20;  Indels   10;  Gaps    6;

Qy        427 AGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTTCCT 486
              || |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        177 AGTACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAGAGGTCCAGTATGCTTCCT 236

Qy        487 ACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAGGATTTG 546
              |||||||||||||||| |||||||||||||||||||||||||| ||||||||||||||||
Db        237 ACCATGGCAGCAGGTCCGAACTCCATACTCTTTGCGATTAACATTGACAACAAGGATTTG 296

Qy        547 AACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAAC 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 AACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAAT 356

Qy        607 GTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCACA 666
              ||||   |||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        357 GTGA---CCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCGTGTTCAGGGAGATCACA 413

Qy        667 GCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCTGTC 726
              |||||||||||| |||||||||||||||||||||||||||||||||||||||||||| |||
Db        414 GCCTCCTCTGCCATCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCCGTC 473

Qy        727 GTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGAAA 786
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        474 GTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGCAA 533

Qy        787 GTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGTTT 846
              |||||||||||| |||||||||||||||||||||||||||||||||||||| |||||
Db        534 GTGACCAACATTATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACATGTTT 593

Qy        847 -AAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCAA 905
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        594 AAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCACTCCTCTGTCAA 653

Qy        906 TATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGCAGCAA 965
              |||||||||||||||||||| |||||||||||| |||||||||||||||||||||||||
Db        654 TATCATGTCCGCCTTTGAAAG-CCCTTCGGGCCAAACTCCAATGGAAGTCAGAGCAGCAA 712

Qy        966 CTC-ATGGAAAATTGATTCCCTGTC--TTTGAACGGGGAGTTTGGGTACCAGAAGCTTGA 1022
              |||  ||||||||||||||||||  ||  |  ||||||| ||||||||||||||||||||
Db        713 CTCGTTGGAAAATTGATTCCCTGTCCTTTGAACAGGGGAGTCTGGGTACCAGAAGCTTGA 772

Qy       1023 TGTGAAAAGCATCGATGATGAAGATGT--GGATGAAAACGAAGATGA 1067
                  |||||||||||||||||||||||||  |||||||||||||||| |
Db        773 TGTGAAAAGCATCGATGATGAAGATGTTGGGATGAAAACGAAGATTA 819


RESULT 15
BQ184171/c
LOCUS       BQ184171                626 bp    mRNA    linear   EST 30-APR-2002
DEFINITION  UI-E-EJ1-ajs-e-23-0-UI.s1 UI-E-EJ1 Homo sapiens cDNA clone
            UI-E-EJ1-ajs-e-23-0-UI 3', mRNA sequence.
ACCESSION   BQ184171
VERSION     BQ184171.1  GI:20359722
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 626)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .626
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-EJ1-ajs-e-23-0-UI"
                     /tissue_type="fetal eyes, lens, eye anterior segment,
                     optic nerve, retina, Retina Foveal and Macular, RPE and
                     Choroid"
                     /dev_stage="fetal and adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-EJ1"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-EJ1 is a subtracted cDNA library constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tags for this library are: fetal eyes,
                     AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,
                     AATGCCGCAT; optic nerve, CCATTAAGTG; retina, CCGCG; Retina
                     Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This
                     library was created for the program, Gene Discovery in the
                     Visual System, supported by National Eye Institute (NEI).
                     TAG_TISSUE=human retina
                     TAG_LIB=UI-E-EJ1
                     TAG_SEQ=CCGCG"
ORIGIN

  Query Match             30.0%;  Score 534.6;  DB 5;  Length 626;
```

```
  Best Local Similarity   99.1%;  Pred. No. 8.3e-133;
  Matches  537;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        245 CAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAGCGCT 304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 CAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAGCGCT 485

Qy        305 TCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGCTGTG 364
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 TNTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGCTGTG 425

Qy        365 CCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGCAGGC 424
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 CCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGCAGGC 365

Qy        425 TGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTTC 484
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 TGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTTC 305

Qy        485 CTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAGGATT 544
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 CTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAGGATT 245

Qy        545 TGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGA 604
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 TGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGA 185

Qy        605 ACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCA 664
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 ACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCA 125

Qy        665 CAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCTG 724
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 CAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCTG 65

Qy        725 TCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGA 784
              ||||||||||||||||||||||||||||||||||||||||| || || ||||| |
Db         64 TCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGAAAAAAAAAAAAAA 5

Qy        785 AA 786
              ||
Db          4 AA 3


Search completed: October 26, 2005, 05:49:54
Job time : 5988 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 25, 2005, 21:11:53 ; Search time 1482 Seconds
(without alignments)
9923.208 Million cell updates/sec

Title: US-10-070-255-4
Perfect score: 1782
Sequence: 1 atgagcggatcacagaacaa.....taaattgctttcattccttg 1782

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9772377 seqs, 4126317084 residues

Total number of hits satisfying chosen parameters: 19544754

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq2:*
25: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
27: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
28: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	275.8	15.5	280	9	US-09-563-817-201	Sequence 201, App
	2	247	13.9	454	24	US-10-450-763-214	Sequence 214, App
c	3	227.8	12.8	542	17	US-10-029-386-8542	Sequence 8542, Ap
c	4	224.8	12.6	284	17	US-10-029-386-22242	Sequence 22242, A
	5	207.6	11.6	463	24	US-10-450-763-28004	Sequence 28004, A
	6	206.8	11.6	465	26	US-11-060-756-3234	Sequence 3234, Ap
	7	206.8	11.6	465	26	US-11-060-756-7506	Sequence 7506, Ap
c	8	143.8	8.1	492	13	US-09-925-065A-383678	Sequence 383678,
c	9	135.4	7.6	474	24	US-10-450-763-9353	Sequence 9353, Ap
	10	128	7.2	500	17	US-10-029-386-8336	Sequence 8336, Ap
	11	124	7.0	124	17	US-10-029-386-22036	Sequence 22036, A
c	12	119	6.7	911	24	US-10-450-763-215	Sequence 215, App
	13	75	4.2	633	13	US-09-925-065A-681221	Sequence 681221,
	14	66.8	3.7	1796	26	US-11-097-143-11288	Sequence 11288, A
c	15	66.8	3.7	3796	26	US-11-097-143-11287	Sequence 11287, A
c	16	63.8	3.6	3931	19	US-10-276-774-948	Sequence 948, App
	17	63	3.5	166536	24	US-10-981-277-35	Sequence 35, Appl
	18	54.4	3.1	813	24	US-10-450-763-21166	Sequence 21166, A
c	19	54.4	3.1	79977	24	US-10-737-082-58	Sequence 58, Appl
c	20	54.4	3.1	79977	24	US-10-765-790-58	Sequence 58, Appl
	21	54	3.0	464	10	US-09-918-995-15177	Sequence 15177, A
c	22	52.8	3.0	263	9	US-09-867-701-9696	Sequence 9696, Ap
	23	52.8	3.0	476	10	US-09-854-867-118	Sequence 118, App
	24	52.8	3.0	476	22	US-10-786-970A-118	Sequence 118, App
	25	50	2.8	371	10	US-09-918-995-24284	Sequence 24284, A
c	26	50	2.8	446	9	US-09-880-107-1518	Sequence 1518, Ap
	27	50	2.8	7093	16	US-10-101-510-527	Sequence 527, App
	28	50	2.8	10126	18	US-10-242-355-1024	Sequence 1024, Ap
	29	49.6	2.8	7140	22	US-10-211-028-3	Sequence 3, Appli
	30	49.6	2.8	90597	22	US-10-211-028-1	Sequence 1, Appli
	31	49.2	2.8	18535	9	US-09-764-878-385	Sequence 385, App
	32	49.2	2.8	18535	15	US-10-079-854-385	Sequence 385, App
	33	48	2.7	2463	14	US-10-027-632-102953	Sequence 102953,
	34	48	2.7	2463	18	US-10-027-632-102953	Sequence 102953,
c	35	48	2.7	2927	24	US-10-450-763-24335	Sequence 24335, A
	36	47.6	2.7	5430	24	US-10-450-763-21169	Sequence 21169, A
	37	47.2	2.6	2227	26	US-11-097-143-14315	Sequence 14315, A
	38	47.2	2.6	4602	26	US-11-097-143-14314	Sequence 14314, A
c	39	46.6	2.6	473	13	US-09-925-065A-844196	Sequence 844196,
c	40	44.4	2.5	505	14	US-10-027-632-276057	Sequence 276057,
c	41	44.4	2.5	505	18	US-10-027-632-276057	Sequence 276057,
c	42	44	2.5	4216	20	US-10-437-963-53652	Sequence 53652, A
c	43	43.8	2.5	1599	20	US-10-437-963-2771	Sequence 2771, Ap
c	44	43.2	2.4	505	14	US-10-027-632-276059	Sequence 276059,
c	45	43.2	2.4	505	18	US-10-027-632-276059	Sequence 276059,

ALIGNMENTS

RESULT 1
US-09-563-817-201
; Sequence 201, Application US/09563817
; Patent No. US20020095031A1
; GENERAL INFORMATION:
; APPLICANT: Nehls, Michael C.
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020095031A1el Human Polynucleotides and the
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: LEX-0021-USA
; CURRENT APPLICATION NUMBER: US/09/563,817
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: US 60/132,343
; PRIOR FILING DATE: 1999-05-04
; NUMBER OF SEQ ID NOS: 1008
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 201
; LENGTH: 280
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(280)
; OTHER INFORMATION: n = A,T,C or G
US-09-563-817-201

Query Match 15.5%; Score 275.8; DB 9; Length 280;

```
  Best Local Similarity   98.9%;  Pred. No. 1.3e-74;
  Matches  277;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1349 GTCCTCTGAGCAGCCTGTTACCTTCTGCCTCAGTGCCAGAGTCCATGACAATTAGTGAAC 1408
              | ||| |||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db             1 GGCCTATGAGCAGCCTGNTACCTTCTGCCTCAGTGCCAGAGTCCATGACAATTAGTGAAC 60

Qy       1409 TGCGCCAGGCCACTGTGGCCATGATGAACAGGAAGGATGAGCTGGAGGAGGAGAACAGAT 1468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            61 TGCGCCAGGCCACTGTGGCCATGATGAACAGGAAGGATGAGCTGGAGGAGGAGAACAGAT 120

Qy       1469 CACTGCGAAACCTGCTCGACGGTGAGATGGAGCACTCAGCCGCGCTCCGGCAAGAGGTGG 1528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           121 CACTGCGAAACCTGCTCGACGGTGAGATGGAGCACTCAGCCGCGCTCCGGCAAGAGGTGG 180

Qy       1529 ACACCTTGAAAAGGAAGGTGGCTGAACAGGAGGAGCGGCAGGGCATGAAGGTCCAGGCGC 1588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           181 ACACCTTGAAAAGGAAGGTGGCTGAACAGGAGGAGCGGCAGGGCATGAAGGTCCAGGCGC 240

Qy       1589 TGGCCAGCTATCTTTGCTATTTTGTGAGGAGATTCTAACC 1628
              |||||||||||||||||||||||||||||||||||||||
Db           241 TGGCCAGCTATCTTTGCTATTTTGTGAGGAGATTCTAACC 280


RESULT 2
US-10-450-763-214
; Sequence 214, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 214
;   LENGTH: 454
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (220)..(387)
;   OTHER INFORMATION: 58% homologous to Homo sapiens Partial seqeunce of the clone
;   OTHER INFORMATION: 9 protein,accession number W42400,Smith-Waterman Score=142.
US-10-450-763-214

  Query Match            13.9%;  Score 247;  DB 24;  Length 454;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-65;
  Matches  247;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 127

Qy         61 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 187

Qy        121 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 247

Qy        181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 307

Qy        241 GAAACAG 247
              |||||||
Db        308 GAAACAG 314


RESULT 3
US-10-029-386-8542/c
; Sequence 8542, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GE
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 8542
;   LENGTH: 542
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO CHR16.3
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
;   OTHER INFORMATION: NT HIT: gi14779172, EVALUE 0.00e+00
;   OTHER INFORMATION: SWISSPROT HIT: O14197, EVALUE 5.90e-01
;   OTHER INFORMATION: EST_HUMAN HIT: BG340557.1, EVALUE 0.00e+00
US-10-029-386-8542

  Query Match            12.8%;  Score 227.8;  DB 17;  Length 542;
  Best Local Similarity  96.1%;  Pred. No. 1.5e-59;
  Matches  245;  Conservative    0;  Mismatches    7;  Indels    3;  Gaps    1;

Qy        497 CAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAGGATTTGAACGGGCAGA 556
              |||||| |||||||||||||||||||||||| ||||||||||||||||||||||||
Db        495 CAGGTCCGAACTCCATACTCTTTGCGATTAACATTGACAACAAGGATTTGAACGGGCAGA 436

Qy        557 GTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAACGTGACCTCCT 616
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||   |||
Db        435 GTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAATGTGA---CCT 379

Qy        617 TGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCACAGCCTCCTCTG 676
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        378 TGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCGTGTTCAGGGAGATCACAGCCTCCTCTG 319

Qy        677 CCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCTGTCGTGTCCAGGA 736
              || |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        318 CCATCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCCGTCGTGTCCAGGA 259

Qy        737 ATGTCAGTGCTGATG 751
              ||||||||||||| ||
Db        258 ATGTCAGTGCTGGTG 244


RESULT 4
US-10-029-386-22242/c
; Sequence 22242, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GE
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 22242
```

```
;   LENGTH: 284
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO CHR16.3
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
;   OTHER INFORMATION: SWISSPROT HIT: O14197, EVALUE 2.10e-01
;   OTHER INFORMATION: NT HIT: gi14779172, EVALUE 0.00e+00
;   OTHER INFORMATION: EST_HUMAN HIT: BF975008.1, EVALUE 0.00e+00
US-10-029-386-22242

  Query Match            12.6%;  Score 224.8;  DB 17;  Length 284;
  Best Local Similarity   96.0%;  Pred. No. 9.2e-59;
  Matches  242;  Conservative    0;  Mismatches    7;  Indels    3;  Gaps    1;

Qy        500 GTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAGGATTTGAACGGGCAGAGTA 559
              ||| |||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db        284 GTCCGAACTCCATACTCTTTGCGATTAACATTGACAACAAGGATTTGAACGGGCAGAGTA 225

Qy        560 AGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAACGTGACCTCCTTGC 619
              |||||||||||||||||||||||||||||||||||||||||||||| ||||   ||||||
Db        224 AGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAATGTGA---CCTTGC 168

Qy        620 TGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCACAGCCTCCTCTGCCG 679
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        167 TGAAGGAGTCCACGCAAGGAGTGAGCAGCGTGTTCAGGGAGATCACAGCCTCCTCTGCCA 108

Qy        680 TCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCTGTCGTGTCCAGGAATG 739
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        107 TCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCCGTCGTGTCCAGGAATG 48

Qy        740 TCAGTGCTGATG 751
              |||||||||  ||
Db         47 TCAGTGCTGGTG 36


RESULT 5
US-10-450-763-28004
; Sequence 28004, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 28004
;   LENGTH: 463
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (318)..(461)
;   OTHER INFORMATION: 66% homologous to Homo sapiens Partial seqeunce of the clone
;   OTHER INFORMATION: 9 protein,accession number W42400,Smith-Waterman Score=124.
US-10-450-763-28004

  Query Match            11.6%;  Score 207.6;  DB 24;  Length 463;
  Best Local Similarity   93.9%;  Pred. No. 2.6e-53;
  Matches  216;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1219 GGCGTGGGCTCCTACAGCCCAGCAGATGCCCCCCTCGGAAGCCTGGAGAACGGGACAGGA 1278
              | | |  || |  | |  ||  ||||||||||||||||||||||||||||||||||||
Db        234 GTCCTCAGCCCTGGCTGAACAAAAGATGCCCCCCTCGGAAGCCTGGAGAACGGGACAGGA 293

Qy       1279 CCAGAGGACCACGTTCTCCCGGATCCTGGACTTCGGTACAGTGTGGAAGCCAGCTCTCCA 1338
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        294 CCAGAGGACCACGTTCTCCCAGATCCTGGACTTCGGTACAGTGTGGAAGCCAGCTCTCCA 353

Qy       1339 GGCCACGGAAGTCCTCTGAGCAGCCTGTTACCTTCTGCCTCAGTGCCAGAGTCCATGACA 1398
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        354 GGCCACGGAAGTCCTCTGAGCAGCCTGTTACCTTCTGCCTCAGTGCCAGAGTCCATGACA 413

Qy       1399 ATTAGTGAACTGCGCCAGGCCACTGTGGCCATGATGAACAGGAAGGATGA 1448
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 ATTAGTGAACTGCGCCAGGCCACTGTGGCCATGATGAACAGGAAGGATGA 463


RESULT 6
US-11-060-756-3234
; Sequence 3234, Application US/11060756
; Publication No. US20050221354A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William Martin
;  TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
;  TITLE OF INVENTION:  Target Genes
;  FILE REFERENCE: AM101083 (031896-042000)
;  CURRENT APPLICATION NUMBER: US/11/060,756
;  CURRENT FILING DATE:  2005-02-18
;  NUMBER OF SEQ ID NOS: 303284
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3234
;   LENGTH: 465
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-060-756-3234

  Query Match            11.6%;  Score 206.8;  DB 26;  Length 465;
  Best Local Similarity   94.5%;  Pred. No. 4.6e-53;
  Matches  225;  Conservative    0;  Mismatches   12;  Indels    1;  Gaps    1;

Qy        274 GAGGTCCTCAACAAGCACGAGCTGC-AGCGCTTCTACTCCCTGCGCCACATCGCCTCAGA 332
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         99 GAGGTCCTCAACAAGCACGAGCTGCAAGCGCTTCTACTCCCTGCGCCACATCGCCTCAGA 158

Qy        333 CGTGGGCCGGGGTCGCGCCTGGCTGCGCTGTGCCCTCAACGAACACTCCCTGGAGCGCTA 392
              ||||||| |||| |||||||||||||||||||||||||||||||||||| ||||||||||
Db        159 CGTGGGCTGGGGCCGCGCCTGGCTGCGCTGTGCCCTCAACGAACACTCCCTGGAGTGCTA 218

Qy        393 CCTGCACATGCTCCTGGCCGACCGCTGCAGGCTGAGCACTTTTTTATGAAGACTGGTCTTT 452
              |||||||||||||||||||||||||||||||| || ||||||||||||||||||||||||
Db        219 CCTGCACATGCTCCTGGCCGACCGCTGCAGGCTCAGTACTTTTTTATGAAGACTGGTCTTT 278

Qy        453 TGTGATGGATGAAGAAAGGTCCAGTATGCTTCCTACCATGGCAGCAGGTCTGAACTCC 510
              |||||||||||||| |||||||||||||||||||||||||||||| |  | |  ||
Db        279 TGTGATGGATGAAGAGAGGTCCAGTATGCTTCCTACCATGGCAGCAGCTGAGGAGGCC 336


RESULT 7
US-11-060-756-7506
; Sequence 7506, Application US/11060756
; Publication No. US20050221354A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William Martin
;  TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
;  TITLE OF INVENTION:  Target Genes
;  FILE REFERENCE: AM101083 (031896-042000)
;  CURRENT APPLICATION NUMBER: US/11/060,756
;  CURRENT FILING DATE:  2005-02-18
;  NUMBER OF SEQ ID NOS: 303284
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7506
;   LENGTH: 465
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-060-756-7506
```

```
  Query Match             11.6%;  Score 206.8;  DB 26;  Length 465;
  Best Local Similarity   94.5%;  Pred. No. 4.6e-53;
  Matches  225;  Conservative    0;  Mismatches   12;  Indels    1;  Gaps    1;

Qy        274 GAGGTCCTCAACAAGCACGAGCTGC-AGCGCTTCTACTCCCTGCGCCACATCGCCTCAGA 332
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         99 GAGGTCCTCAACAAGCACGAGCTGCAAGCGCTTCTACTCCCTGCGCCACATCGCCTCAGA 158

Qy        333 CGTGGGCCGGGGTCGCGCCTGGCTGCGCTGTGCCCTCAACGAACACTCCCTGGAGCGCTA 392
              ||||||| |||| ||||||||||||||||||||||||||||||||||||||||| ||||
Db        159 CGTGGGCTGGGGCCGCGCCTGGCTGCGCTGTGCCCTCAACGAACACTCCCTGGAGTGCTA 218

Qy        393 CCTGCACATGCTCCTGGCCGACCGCTGCAGGCTGAGCACTTTTTATGAAGACTGGTCTTT 452
              |||||||||||||||||||||||||||||||| || ||||||||||||||||||||||
Db        219 CCTGCACATGCTCCTGGCCGACCGCTGCAGGCTCAGTACTTTTTATGAAGACTGGTCTTT 278

Qy        453 TGTGATGGATGAAGAAAGGTCCAGTATGCTTCCTACCATGGCAGCAGGTCTGAACTCC 510
              ||||||||||||||| |||||||||||||||||||||||||||| | | | ||
Db        279 TGTGATGGATGAAGAGAGGTCCAGTATGCTTCCTACCATGGCAGCAGCTGAGGAGGCC 336


RESULT 8
US-09-925-065A-383678/c
; Sequence 383678, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 383678
;   LENGTH: 492
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-383678

  Query Match             8.1%;  Score 143.8;  DB 13;  Length 492;
  Best Local Similarity   98.1%;  Pred. No. 1.9e-33;
  Matches  156;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       1589 TGGCCAGCTATCTTTGCTATTTTGTGAGGAGATTCTAACCCCACGTGAGAACCATGTGGT 1648
              | | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 TTGGCAGCTATCTTTGCTATTTTGTGAGGAGATTCTAACCCCACGTGAGAACCATGTGGT 99

Qy       1649 GGAGAAATGGAGGGAGAGAGAAATCCAACAGTTCCTGATAGTCTCATTTGAGCTCCTGGA 1708
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db          98 GGAGAAATGGAGGGAGAGAGAAAT-CAACAGTTCCTGATAGTCTCATTTGAGCTCCTGGA 40

Qy       1709 TCCAGTCTTTCCTGAAGCTGTGTTTCCTCTGGACTTTTC 1747
              |||||||||||||||||||||||||||||||||||||||
Db          39 TCCAGTCTTTCCTGAAGCTGTGTTTCCTCTGGACTTTTC 1


RESULT 9
US-10-450-763-9353/c
; Sequence 9353, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 9353
;   LENGTH: 474
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (369)..(133)
;   OTHER INFORMATION: 67% homologous to Homo sapiens Partial seqeunce of the clone
;   OTHER INFORMATION: 9 protein,accession number W42400,Smith-Waterman Score=227.
US-10-450-763-9353

  Query Match             7.6%;  Score 135.4;  DB 24;  Length 474;
  Best Local Similarity   72.6%;  Pred. No. 7.5e-31;
  Matches  175;  Conservative    0;  Mismatches   66;  Indels    0;  Gaps    0;

Qy         81 CTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGCAGGGTCACCTGTCTGTGTGC 140
              || |||   ||   | ||       || ||   |      ||||||||||||||||||||
Db        367 CTCTGGCCAGAAGCAAGAAGAAAGTTCTGACCACTGGCAGAGGGTCACCTGTCTGTGTGC 308

Qy        141 CCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGTCGAGGATTGGCACTCACAGC 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGTCGAGGATTGGCACTCACAGC 248

Qy        201 GGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACCGAAACAGAGCCCGTGTTCTG 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||      || |
Db        247 GGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACCGAAACAGGTGAACTGCGCCA 188

Qy        261 GTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAGCGCTTCTACTCCCTGCGCCA 320
              | |   |||     | |   ||  ||| | |||||| ||      ||  |||| | |
Db        187 GGCCACTGTGGCCATGATGAACAGGAAGGATGAGCTGGAGGAGGAGAACAGCCTGGGTGA 128

Qy        321 C 321
              |
Db        127 C 127


RESULT 10
US-10-029-386-8336
; Sequence 8336, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GE
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 8336
;   LENGTH: 500
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO CHR16.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
```

```
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
;   OTHER INFORMATION: SWISSPROT HIT: Q9Z883, EVALUE 1.20e+00
;   OTHER INFORMATION: NT HIT: gi15315787, EVALUE 2.00e-65
;   OTHER INFORMATION: EST_HUMAN HIT: AL135642.1, EVALUE 2.00e-65
US-10-029-386-8336

  Query Match             7.2%;  Score 128;  DB 17;  Length 500;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-28;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        120 CAGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAG 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 CAGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAG 400

Qy        180 TCGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAAC 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 TCGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAAC 460

Qy        240 CGAAACAG 247
              ||||||||
Db        461 CGAAACAG 468


RESULT 11
US-10-029-386-22036
; Sequence 22036, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 22036
;   LENGTH: 124
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO CHR16.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
;   OTHER INFORMATION: NT HIT: gi15315787, EVALUE 9.00e-64
;   OTHER INFORMATION: EST_HUMAN HIT: AL135642.1, EVALUE 1.00e-63
;   OTHER INFORMATION: SWISSPROT HIT: Q9SXU1, EVALUE 6.00e+00
US-10-029-386-22036

  Query Match             7.0%;  Score 124;  DB 17;  Length 124;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-27;
  Matches  124;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        124 GTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGTCGA 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGTCGA 60

Qy        184 GGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACCGAA 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACCGAA 120

Qy        244 ACAG 247
              ||||
Db        121 ACAG 124


RESULT 12
US-10-450-763-215/c
; Sequence 215, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 215
;   LENGTH: 911
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (173)..(57)
;   OTHER INFORMATION: 100% homologous to Homo sapiens Partial seqeunce of the clone
;   OTHER INFORMATION: 9 protein,accession number W42400,Smith-Waterman Score=193.
US-10-450-763-215

  Query Match             6.7%;  Score 119;  DB 24;  Length 911;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-25;
  Matches  119;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        129 CTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGTCGAGGATT 188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 CTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGTCGAGGATT 115

Qy        189 GGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACCGAAACAG 247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        114 GGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACCGAAACAG 56


RESULT 13
US-09-925-065A-681221
; Sequence 681221, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 681221
;   LENGTH: 633
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-681221

  Query Match             4.2%;  Score 75;  DB 13;  Length 633;
  Best Local Similarity   100.0%;  Pred. No. 5.2e-12;
  Matches   75;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy        427 AGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTTCCT 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 AGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTTCCT 543

Qy        487 ACCATGGCAGCAGGT 501
              |||||||||||||||
Db        544 ACCATGGCAGCAGGT 558
```

RESULT 14
US-11-097-143-11288
; Sequence 11288, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11288
; LENGTH: 1796
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-11288

Query Match 3.7%; Score 66.8; DB 26; Length 1796;
Best Local Similarity 49.4%; Pred. No. 3.1e-09;
Matches 173; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

```
Qy        233 GCAAAACCGAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACG 292
              ||| | | | |  ||  ||   |||||| | | |    ||  |    | |    ||||
Db        396 GCAGCAACAACAATGACGCCACTTTCTGGGAATTCTGCCAGACGCATTTGACGCCCCACG 455

Qy        293 AGCTGCAGCGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCT 352
              | |  ||| | |        |||    || |||    |  | ||||| | || || ||||
Db        456 AACGTCAGAGGTATATGGACCTTAAGCAGATCTGGACGAATGTGGGCAGAGGACGTGCCT 515

Qy        353 GGCTGCGCTGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCG 412
                | |||    | || |||||  |    ||| |  || |  ||| ||      ||| ||
Db        516 TCATACGCGCCACACTGAACGAGAAGCGACTGCATAGCCATGTGCTCACCTGGCTGAGCG 575

Qy        413 ACCGCTGCAGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGT 472
              ||        ||||       |||||   |   |||||  |  |  |  | || ||   |
Db        576 ACGAGGAGCAGCTGCATCGATTTTACACACCTTGGTCCCTGCTACTCAACGATGAGGCGG 635

Qy        473 CCAGTATGCTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCG 532
              |||  | |||  ||   || | ||    ||  |    ||||  || ||  | ||  | |
Db        636 CCAAGAAGCTGCCAGAGATCGTAGACTCCCTCAGTGATGTACTGTTCGCCCTCAATGTGG 695

Qy        533 ACAACAAGGATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGAC 582
              ||| ||  ||  | || |  |  || |  |   | || |  ||| | | |
```

```
Db        696 ACACCACTGAGCTTAATGCACCCAGGAGATCGACGCCAAGTGTTGCCGTC 745
```

RESULT 15
US-11-097-143-11287/c
; Sequence 11287, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11287
; LENGTH: 3796
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-11287

Query Match 3.7%; Score 66.8; DB 26; Length 3796;
Best Local Similarity 49.4%; Pred. No. 4.5e-09;
Matches 173; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

```
Qy        233 GCAAAACCGAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACG 292
              ||| | | | |  ||  ||   |||||| | | |    ||  |    | |    ||||
Db       2401 GCAGCAACAACAATGACGCCACTTTCTGGGAATTCTGCCAGACGCATTTGACGCCCCACG 2342

Qy        293 AGCTGCAGCGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCT 352
              | |  ||| | |        |||    || |||    |  | ||||| | || || ||||
Db       2341 AACGTCAGAGGTATATGGACCTTAAGCAGATCTGGACGAATGTGGGCAGAGGACGTGCCT 2282

Qy        353 GGCTGCGCTGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCG 412
                | |||    | || |||||  |    ||| |  || |  ||| ||      ||| ||
Db       2281 TCATACGCGCCACACTGAACGAGAAGCGACTGCATAGCCATGTGCTCACCTGGCTGAGCG 2222

Qy        413 ACCGCTGCAGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGT 472
              ||        ||||       |||||   |   |||||  |  |  |  | || ||   |
Db       2221 ACGAGGAGCAGCTGCATCGATTTTACACACCTTGGTCCCTGCTACTCAACGATGAGGCGG 2162

Qy        473 CCAGTATGCTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCG 532
              |||  | |||  ||   || | ||    ||  |    ||||  || ||  | ||  | |
Db       2161 CCAAGAAGCTGCCAGAGATCGTAGACTCCCTCAGTGATGTACTGTTCGCCCTCAATGTGG 2102

Qy        533 ACAACAAGGATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGAC 582
              ||| ||  ||  | || |  |  || |  |   | || |  ||| | | |
Db       2101 ACACCACTGAGCTTAATGCACCCAGGAGATCGACGCCAAGTGTTGCCGTC 2052
```

Search completed: October 26, 2005, 01:43:40
Job time : 1484 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 25, 2005, 18:28:14 ; Search time 286 Seconds
(without alignments)
10195.263 Million cell updates/sec

Title: US-10-070-255-4
Perfect score: 1782
Sequence: 1 atgagcggatcacagaacaa.....taaattgctttcattccttg 1782

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	133.4	7.5	190	4	US-09-513-999C-16002	Sequence 16002, A
	2	66.8	3.7	811	4	US-09-270-767-12402	Sequence 12402, A
c	3	56.8	3.2	7218	1	US-08-232-463-14	Sequence 14, Appl
c	4	54.4	3.1	36731	4	US-09-949-016-13770	Sequence 13770, A
c	5	54	3.0	601	4	US-09-949-016-69138	Sequence 69138, A
	6	52.8	3.0	476	4	US-09-573-080A-118	Sequence 118, App
	7	52.2	2.9	961	4	US-09-976-594-221	Sequence 221, App
	8	49.6	2.8	9811	4	US-09-949-016-12759	Sequence 12759, A
	9	47	2.6	100	4	US-09-513-999C-16212	Sequence 16212, A
c	10	46.8	2.6	60276	4	US-09-949-016-15004	Sequence 15004, A
c	11	46.8	2.6	60338	4	US-09-949-016-15694	Sequence 15694, A
c	12	46.8	2.6	94133	4	US-09-949-016-11901	Sequence 11901, A
c	13	46.8	2.6	94133	4	US-09-949-016-12713	Sequence 12713, A
c	14	46.8	2.6	94135	4	US-09-949-016-15934	Sequence 15934, A
c	15	46.8	2.6	94135	4	US-09-949-016-15935	Sequence 15935, A
c	16	46.8	2.6	94135	4	US-09-949-016-15936	Sequence 15936, A
c	17	46.8	2.6	94135	4	US-09-949-016-15937	Sequence 15937, A
	18	44.8	2.5	474	4	US-09-621-976-18033	Sequence 18033, A
	19	43.2	2.4	1002	4	US-09-902-540-8383	Sequence 8383, Ap
c	20	43.2	2.4	7035	4	US-09-902-540-878	Sequence 878, App
c	21	43.2	2.4	23802	4	US-09-949-016-12107	Sequence 12107, A
c	22	43.2	2.4	23803	4	US-09-949-016-15878	Sequence 15878, A
	23	43	2.4	1304	4	US-09-902-540-2596	Sequence 2596, Ap
	24	43	2.4	16584	4	US-09-902-540-1119	Sequence 1119, Ap
	25	42	2.4	23417	4	US-09-902-540-1207	Sequence 1207, Ap
c	26	41.4	2.3	505	4	US-09-621-976-15639	Sequence 15639, A
	27	41.2	2.3	1926	3	US-09-249-585A-2	Sequence 2, Appli
	28	41.2	2.3	1926	4	US-09-410-399-3	Sequence 3, Appli
	29	41.2	2.3	2580	3	US-09-050-863-2	Sequence 2, Appli
	30	41.2	2.3	2580	3	US-09-359-081-2	Sequence 2, Appli
c	31	41.2	2.3	5452	2	US-09-130-114-1	Sequence 1, Appli
c	32	41.2	2.3	8705	4	US-09-647-344A-14	Sequence 14, Appl
	33	41.2	2.3	9600	3	US-08-910-647-1	Sequence 1, Appli
	34	41.2	2.3	9600	3	US-09-620-925-1	Sequence 1, Appli
	35	41.2	2.3	10596	1	US-07-884-811-15	Sequence 15, Appl
	36	41.2	2.3	10596	1	US-07-885-971-15	Sequence 15, Appl
	37	41.2	2.3	10596	1	US-08-087-783A-15	Sequence 15, Appl
	38	41.2	2.3	10596	1	US-08-194-088B-15	Sequence 15, Appl
	39	41.2	2.3	10596	2	US-08-194-087-15	Sequence 15, Appl
	40	41.2	2.3	10596	5	PCT-US93-04648-15	Sequence 15, Appl
c	41	41.2	2.3	16080	4	US-09-724-566A-48	Sequence 48, Appl
c	42	41.2	2.3	16080	4	US-09-471-669A-48	Sequence 48, Appl
	43	41	2.3	1275	4	US-09-252-991A-251	Sequence 251, App
c	44	41	2.3	1785	4	US-09-252-991A-281	Sequence 281, App
	45	40.6	2.3	1132	3	US-09-370-838-256	Sequence 256, App

ALIGNMENTS

RESULT 1
US-09-513-999C-16002
; Sequence 16002, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 16002
; LENGTH: 190
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-16002

Query Match 7.5%; Score 133.4; DB 4; Length 190;
Best Local Similarity 99.3%; Pred. No. 8.4e-26;
Matches 134; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1464 CAGATCACTGCGAAACCTGCTCGACGGTGAGATGGAGCACTCAGCCGCGCTCCGGCAAGA 1523
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 CAGATCACTGCGAAACCTGCTCGACGGTGAGATGGAGCACTCAGCCGCGCTCCGGCAAGA 78

Qy       1524 GGTGGACACCTTGAAAAGGAAGGTGGCTGAACAGGAGGAGCGGCAGGGCATGAAGGTCCA 1583
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 GGTGGACACCTTGAAAAGGAAGGTGGCTGAACAGGAGGAGCGGCAGGGCATGAAGGTCCA 138

Qy       1584 GGCGCTGGCCAGCTA 1598
             ||||||||||||| ||
Db        139 GGCGCTGGCCAGGTA 153
```

RESULT 2
US-09-270-767-12402
; Sequence 12402, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17

```
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12402
;   LENGTH: 811
;   TYPE: DNA
;   ORGANISM: Drosophila melanogaster
US-09-270-767-12402

  Query Match             3.7%;  Score 66.8;  DB 4;  Length 811;
  Best Local Similarity   49.4%;  Pred. No. 1e-07;
  Matches  173;  Conservative    0;  Mismatches  177;  Indels    0;  Gaps    0;

Qy        233 GCAAAACCGAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACG 292
              ||| | | | |  ||  ||   |||||| | | |    ||  |    | |     ||||
Db        396 GCAGCAACAACAATGACGCCACTTTCTGGGAATTCTGCCAGACGCATTTGACGCCCCACG 455

Qy        293 AGCTGCAGCGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCT 352
              | |  ||| | |       |||    || |||    |  | |||||| | || || ||||
Db        456 AACGTCAGAGGTATATGGACCTTAAGCAGATCTGGACGAATGTGGGCAGAGGACGTGCCT 515

Qy        353 GGCTGCGCTGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCG 412
                | |||    | || |||||  |    ||| |  || |  ||| ||      |||  ||
Db        516 TCATACGCGCCACACTGAACGAGAAGCGACTGCATAGCCATGTGCTCACCTGGCTGAGCG 575

Qy        413 ACCGCTGCAGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGT 472
              ||        ||||      |||||    |   |||| |  |  |  | || ||   |
Db        576 ACGAGGAGCAGCTGCATCGATTTTACACACCTTGGTCCCTGCTACTCAACGATGAGGCGG 635

Qy        473 CCAGTATGCTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCG 532
              |||  | ||| ||    || | ||     || |       |||| || ||  | ||  | |
Db        636 CCAAGAAGCTGCCAGAGATCGTAGACTCCCTCAGTGATGTACTGTTCGCCCTCAATGTGG 695

Qy        533 ACAACAAGGATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGAC 582
              ||| ||  ||  | || |  |  || |  |    | || |  ||| | | |
Db        696 ACACCACTGAGCTTAATGCACCCAGGAGATCGACGCCAAGTGTTGCCGTC 745


RESULT 3
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14

  Query Match             3.2%;  Score 56.8;  DB 1;  Length 7218;
  Best Local Similarity   3.3%;  Pred. No. 0.00013;
  Matches   13;  Conservative  225;  Mismatches  152;  Indels    0;  Gaps    0;

Qy        734 GGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGAAAGTGACCA 793
              || | :  :: :  :: :  ::: : ::::::::: :::::::::::::::::: ::  :
Db       1439 GGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1380

Qy        794 ACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGTTTAAAAAGA 853
              : : ::    :   :: :: :::::::: ::: ::::    ::::: : :   :::::::
Db       1379 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1320

Qy        854 CACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCAATATCATGT 913
              :   :::: :::::::::   :::::: ::    ::  :        :  :: :  : :
Db       1319 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1260

Qy        914 CCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGCAGCAACTCATGGA 973
                :     :::::       :::   ::    :: :::::  :::: :: ::   : :::
Db       1259 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1200

Qy        974 AAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTTGATGTGAAAAGCA 1033
              ::: ::       :     ::: :::::   ::: :  :::::   :: : :::::: :
Db       1199 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1140

Qy       1034 TCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCATCAGGAAGGA 1093
              :: :: ::::: : ::: ::::: ::::: :: : : : :::::   :  ::::::::
Db       1139 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1080

Qy       1094 AGCACAGGGGCCACTCGGAGTCGCCCGAGA 1123
              :: : :::::  : ||| |  | |||  ||
Db       1079 RRRRRRRRRRRRATCGCAAGCTCCCTCGA 1050


RESULT 4
US-09-949-016-13770/c
; Sequence 13770, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13770
;   LENGTH: 36731
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13770

  Query Match             3.1%;  Score 54.4;  DB 4;  Length 36731;
  Best Local Similarity   64.1%;  Pred. No. 0.0012;
```

```
  Matches   82;  Conservative    0;  Mismatches   46;  Indels    0;  Gaps    0;

Qy       1654 AATGGAGGGAGAGAGAAATCCAACAGTTCCTGATAGTCTCATTTGAGCTCCTGGATCCAG 1713
              ||  |  | |   |||     ||   ||||||||   ||||||  ||   ||||||||||
Db       6869 AAAAGTTGAACCAAGACGGAGAAAGATTCCTGATGACCTCATTAGAATACCTGGATCCAG 6810

Qy       1714 TCTTTCCTGAAGCTGTGTTTCCTCTGGACTTTTCATGTATGTGAGCCAATAAATTGCTTT 1773
              ||| ||||||       | ||||||| || | |   ||||||||||| ||| || |
Db       6809 CCTTGCCTGAATGCAGTATACCTCTGGATTTCTTAGTCATGTGAGCCAATGAATCGCCTG 6750

Qy       1774 CATTCCTT 1781
                ||  |||
Db       6749 TTTTGCTT 6742


RESULT 5
US-09-949-016-69138/c
; Sequence 69138, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 69138
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-69138

  Query Match             3.0%;  Score 54;  DB 4;  Length 601;
  Best Local Similarity   63.3%;  Pred. No. 0.00024;
  Matches   81;  Conservative    1;  Mismatches   46;  Indels    0;  Gaps    0;

Qy       1654 AATGGAGGGAGAGAGAAATCCAACAGTTCCTGATAGTCTCATTTGAGCTCCTGGATCCAG 1713
              ||  |  | |   |||     ||   ||||||||   ||||||  ||   ||||||||||
Db        414 AAAAGTTGAACCAAGACGGAGAAAGATTCCTGATGACCTCATTAGAATACCTGGATCCAG 355

Qy       1714 TCTTTCCTGAAGCTGTGTTTCCTCTGGACTTTTCATGTATGTGAGCCAATAAATTGCTTT 1773
              ||| ||||||       | ||||||| || | |   ||||||||||| ||: || |
Db        354 CCTTGCCTGAATGCAGTATACCTCTGGATTTCTTAGTCATGTGAGCCAATGAAKCGCCTG 295

Qy       1774 CATTCCTT 1781
                ||  |||
Db        294 TTTTGCTT 287


RESULT 6
US-09-573-080A-118
; Sequence 118, Application US/09573080A
; Patent No. 6828097
; GENERAL INFORMATION:
;  APPLICANT:  JOAN, KNOLL
;  APPLICANT:  ROGAN, PETER
;  TITLE OF INVENTION:   SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATI
;  FILE REFERENCE:   30307
;  CURRENT APPLICATION NUMBER: US/09/573,080A
;  CURRENT FILING DATE:  2000-05-16
;  NUMBER OF SEQ ID NOS:   479
;  SOFTWARE:   PatentIn version 3.0
; SEQ ID NO 118
;   LENGTH: 476
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: repeat_region
;   LOCATION: (1)..(476)
;  PUBLICATION INFORMATION:
;  PUBLICATION INFORMATION:
;   AUTHORS: Jurka, J; Walichiewicz, J; Milosavljevic, A
;   TITLE: Prototypic sequences for human repetitive DNA
;   JOURNAL: Journal of Molecular Evolution
;   VOLUME: 35
;   ISSUE: 4
;   PAGES: 286-291
;   DATE: 1992-10-__
;   DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
;   DATABASE ENTRY DATE: ____-__-__
;   DATABASE ENTRY DATE: 1996-01-26
US-09-573-080A-118

  Query Match             3.0%;  Score 52.8;  DB 4;  Length 476;
  Best Local Similarity   69.8%;  Pred. No. 0.00045;
  Matches   88;  Conservative    0;  Mismatches   32;  Indels    6;  Gaps    1;

Qy       1662 GAGAGAGAAATCCAACAGTTCCTGATAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCCT 1721
              ||||  || |   |    ||||||||    |||||||||| |||||||||| | | |||
Db        313 GAGATGGAGAGAGACAGATTCCTGATGACATCATTTGAGCACCTGGATCCAGCCATGCCT 372

Qy       1722 GAAG------CTGTGTTTCCTCTGGACTTTTCATGTATGTGAGCCAATAAATTGCTTTCA 1775
              ||||      || || | || ||||||||||||  ||  ||||||||||||||   ||
Db        373 GAAGCCAGATCTATGCTACCCCTGGACTTTTCAGTTACATGAGCCAATAAATTCTCTTTT 432

Qy       1776 TTCCTT 1781
                ||  |||
Db        433 TTGCTT 438


RESULT 7
US-09-976-594-221
; Sequence 221, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
;  APPLICANT: Furness, Michael
;  APPLICANT:  Buchbinder, Jenny
;  TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
;  FILE REFERENCE: PA-0041 US
;  CURRENT APPLICATION NUMBER: US/09/976,594
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: 60/240,409
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: 1143
;  SOFTWARE: PERL Program
; SEQ ID NO 221
;   LENGTH: 961
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6673549 330927.1
US-09-976-594-221

  Query Match             2.9%;  Score 52.2;  DB 4;  Length 961;
  Best Local Similarity   71.1%;  Pred. No. 0.0009;
  Matches   69;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy       1679 GTTCCTGATAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCCTGAAGCTGTGTTTCCTCT 1738
              || | || |    || |||||||  || || | ||  || |||||||| | | ||||
Db        861 GTCCTTGGTGACATCTTTTGAGCCACTAGACCAAGCTTTACCTGAAGCAGAGCTACCTCA 920

Qy       1739 GGACTTTTCATGTATGTGAGCCAATAAATTGCTTTCA 1775
              | ||||||||  ||||||||||||||||    || |||
Db        921 GAACTTTTCAGCTATGTGAGCCAATAAACATCTGTCA 957


RESULT 8
```

```
US-09-949-016-12759
; Sequence 12759, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12759
;   LENGTH: 9811
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12759

  Query Match             2.8%;  Score 49.6;  DB 4;  Length 9811;
  Best Local Similarity   61.7%;  Pred. No. 0.013;
  Matches   79;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy       1654 AATGGAGGGAGAGAGAAATCCAACAGTTCCTGATAGTCTCATTTGAGCTCCTGGATCCAG 1713
              ||   || ||| |||  |   | || ||| | ||    || ||||||  |||  | | |
Db       3752 AAGAAAGAGAGTGAGTGAGTGACCATTTCTTTATTCCATCCTTTGAGGCCCTACACCTAA 3811

Qy       1714 TCTTTCCTGAAGCTGTGTTTCCTCTGGACTTTTCATGTATGTGAGCCAATAAATTGCTTT 1773
              || ||| ||||||| |    |  |  | |||||||  ||  ||||||||| |||| | ||
Db       3812 TCATTCATGAAGCTATACCACTCCCAGCCTTTTCAATTACTTGAGCCAATCAATTTCCTT 3871

Qy       1774 CATTCCTT 1781
                | | ||
Db       3872 TAATAGTT 3879


RESULT 9
US-09-513-999C-16212
; Sequence 16212, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 16212
;   LENGTH: 100
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 13
;   OTHER INFORMATION: m=a or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 14
;   OTHER INFORMATION: m=a or c
US-09-513-999C-16212

  Query Match             2.6%;  Score 47;  DB 4;  Length 100;
  Best Local Similarity   100.0%;  Pred. No. 0.0078;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCG 47
              |||||||||||||||||||||||||||||||||||||||||||||||
Db         53 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCG 99


RESULT 10
US-09-949-016-15004/c
; Sequence 15004, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15004
;   LENGTH: 60276
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(60276)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15004

  Query Match             2.6%;  Score 46.8;  DB 4;  Length 60276;
  Best Local Similarity   73.2%;  Pred. No. 0.17;
  Matches   60;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       1692 TCATTTGAGCTCCTGGATCCAGTCTTTCCTGAAGCTGTGTTTCCTCTGGACTTTTCATGT 1751
              ||  |   ||| |||||||||||| | | ||||||||  |  | || |||||||| |||  |
Db      10119 TCTTAGAAGCACCTGGATCCAGCCCTGCCTGAAGCCATCCTACCCCTGGACTTCTCAGTT 10060

Qy       1752 ATGTGAGCCAATAAATTGCTTT 1773
                |||||||||| || |  ||||
Db      10059 GTGTGAGCCAAGAATTGTCTTT 10038


RESULT 11
US-09-949-016-15694/c
; Sequence 15694, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15694
;   LENGTH: 60338
;   TYPE: DNA
;   ORGANISM: Human
```

```
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(60338)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15694

  Query Match             2.6%;  Score 46.8;  DB 4;  Length 60338;
  Best Local Similarity   73.2%;  Pred. No. 0.17;
  Matches   60;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       1692 TCATTTGAGCTCCTGGATCCAGTCTTTCCTGAAGCTGTGTTTCCTCTGGACTTTTCATGT 1751
              ||  |   ||| |||||||||||| | | ||||||||  |  | || ||||||||| |||  |
Db      10181 TCTTAGAAGCACCTGGATCCAGCCCTGCCTGAAGCCATCCTACCCCTGGACTTCTCAGTT 10122

Qy       1752 ATGTGAGCCAATAAATTGCTTT 1773
              |||||||||| || |  ||||
Db      10121 GTGTGAGCCAAGAATTGTCTTT 10100


RESULT 12
US-09-949-016-11901/c
; Sequence 11901, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11901
;   LENGTH: 94133
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-11901

  Query Match             2.6%;  Score 46.8;  DB 4;  Length 94133;
  Best Local Similarity   67.3%;  Pred. No. 0.21;
  Matches   66;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy       1678 AGTTCCTGATAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCCTGAAGCTGTGTTTCCTC 1737
              ||  |||||  |       ||| |||   ||||  |||   | ||||||||  || | ||||
Db      40227 AGAGCCTGAGAACACAGTTTAAGCATTTGGAGTCAGCTATGCCTGAAGCCATGCTCCCTC 40168

Qy       1738 TGGACTTTTCATGTATGTGAGCCAATAAATTGCTTTCA 1775
              | ||||||||||  ||  |||||||||||||  |  ||| |
Db      40167 TAGACTTTTCAGTTACATGAGCCAATAAAATATTTTTA 40130


RESULT 13
US-09-949-016-12713/c
; Sequence 12713, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12713
;   LENGTH: 94133
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12713

  Query Match             2.6%;  Score 46.8;  DB 4;  Length 94133;
  Best Local Similarity   67.3%;  Pred. No. 0.21;
  Matches   66;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy       1678 AGTTCCTGATAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCCTGAAGCTGTGTTTCCTC 1737
              ||  |||||  |       ||| |||   ||||  |||   | ||||||||  || | ||||
Db      40227 AGAGCCTGAGAACACAGTTTAAGCATTTGGAGTCAGCTATGCCTGAAGCCATGCTCCCTC 40168

Qy       1738 TGGACTTTTCATGTATGTGAGCCAATAAATTGCTTTCA 1775
              | ||||||||||  ||  |||||||||||||  |  ||| |
Db      40167 TAGACTTTTCAGTTACATGAGCCAATAAAATATTTTTA 40130


RESULT 14
US-09-949-016-15934/c
; Sequence 15934, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15934
;   LENGTH: 94135
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-15934

  Query Match             2.6%;  Score 46.8;  DB 4;  Length 94135;
  Best Local Similarity   67.3%;  Pred. No. 0.21;
  Matches   66;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy       1678 AGTTCCTGATAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCCTGAAGCTGTGTTTCCTC 1737
              ||  |||||  |       ||| |||   ||||  |||   | ||||||||  || | ||||
Db      40227 AGAGCCTGAGAACACAGTTTAAGCATTTGGAGTCAGCTATGCCTGAAGCCATGCTCCCTC 40168

Qy       1738 TGGACTTTTCATGTATGTGAGCCAATAAATTGCTTTCA 1775
              | ||||||||||  ||  |||||||||||||  |  ||| |
Db      40167 TAGACTTTTCAGTTACATGAGCCAATAAAATATTTTTA 40130


RESULT 15
US-09-949-016-15935/c
; Sequence 15935, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
```

; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15935
; LENGTH: 94135
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15935

Query Match 2.6%; Score 46.8; DB 4; Length 94135;
Best Local Similarity 67.3%; Pred. No. 0.21;
Matches 66; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

```
Qy       1678 AGTTCCTGATAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCCTGAAGCTGTGTTTCCTC 1737
              ||  ||||| |       ||| |||   ||||  |||   | ||||||||  || | ||||
Db      40227 AGAGCCTGAGAACACAGTTTAAGCATTTGGAGTCAGCTATGCCTGAAGCCATGCTCCCTC 40168

Qy       1738 TGGACTTTTCATGTATGTGAGCCAATAAATTGCTTTCA 1775
              | |||||||||  ||  ||||||||||||| |  ||| |
Db      40167 TAGACTTTTCAGTTACATGAGCCAATAAAATATTTTTA 40130
```

Search completed: October 25, 2005, 21:20:34
Job time : 289 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 25, 2005, 21:16:28 ; Search time 981 Seconds
(without alignments)
10753.292 Million cell updates/sec

Title: US-10-070-255-4
Perfect score: 1782
Sequence: 1 atgagcggatcacagaacaa.....taaattgctttcattccttg 1782

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		1782	100.0	1782	4	AAD04333	Aad04333 Human TNF
2		1742	97.8	1906	2	AAV03323	Aav03323 5' end of
3		1595	89.5	2873	4	AAD04335	Aad04335 Human TNF
4		1595	89.5	3139	4	AAD04334	Aad04334 Human TNF
5		1129.8	63.4	2248	4	AAH15665	Aah15665 Human cDN
6		816.8	45.8	6045	4	AAH18235	Aah18235 Human cDN
7		639.8	35.9	730	4	AAH04973	Aah04973 Human cDN
8		364.2	20.4	33147	4	AAK67282	Aak67282 Human imm
9		275.8	15.5	280	6	ABS72233	Abs72233 Human gen
10		247	13.9	454	5	AAS64410	Aas64410 DNA encod
c 11		227.8	12.8	542	12	ACH75347	Ach75347 Human gen
c 12		224.8	12.6	284	12	ACH89047	Ach89047 Human gen
13		207.6	11.6	463	5	AAS92200	Aas92200 DNA encod
c 14		135.4	7.6	474	5	AAS73549	Aas73549 DNA encod
15		133.4	7.5	190	3	AAC11927	Aac11927 Human sec
16		128	7.2	500	12	ACH75141	Ach75141 Human gen
17		124	7.0	124	12	ACH88841	Ach88841 Human gen
c 18		119	6.7	911	5	AAS64411	Aas64411 DNA encod
19		66.8	3.7	1796	4	ABL09365	Abl09365 Drosophil
c 20		66.8	3.7	3796	4	ABL09364	Abl09364 Drosophil
c 21		63.8	3.6	3931	4	ABA09172	Aba09172 Human BG3
c 22		59.6	3.3	214520	10	ADL13471	Adl13471 Osteoarth
23		59	3.3	3927	10	ADF81729	Adf81729 Leukaemia
24		58.4	3.3	595	13	ADQ53383	Adq53383 Novel can
25		54.4	3.1	813	5	AAS85362	Aas85362 DNA encod
26		54	3.0	464	9	ACH27965	Ach27965 Human adu
c 27		52.8	3.0	263	6	ABL86718	Abl86718 Human ova
28		52.8	3.0	476	7	ADS31085	Ads31085 Human gen
29		52.2	2.9	729	10	ADE07225	Ade07225 Novel cod
30		52.2	2.9	941	5	ABA21208	Aba21208 Human ner
31		52.2	2.9	941	5	ABA21210	Aba21210 Human ner
32		52.2	2.9	961	12	ADL12492	Adl12492 Human ste
33		52.2	2.9	3916	13	ADR07577	Adr07577 Full leng
34		51.8	2.9	617	5	ABA14117	Aba14117 Human ner
35		50.2	2.8	2000	8	ADA71938	Ada71938 Rice gene
36		50	2.8	371	9	ACH37072	Ach37072 Human end
c 37		50	2.8	446	6	ABN95020	Abn95020 Gene #151
c 38		50	2.8	769	4	AAI95747	Aai95747 Human neu
39		50	2.8	5684	3	AAC59953	Aac59953 Human sec
40		50	2.8	5684	8	ABZ73290	Abz73290 Secreted
41		50	2.8	5684	10	ABZ66903	Abz66903 Human sec
42		50	2.8	7093	6	ABZ35416	Abz35416 Human gen
43		50	2.8	10126	4	AAI99260	Aai99260 Human exc
44		50	2.8	10126	5	AAI63610	Aai63610 Human kid
45		49.6	2.8	90597	10	ADJ72363	Adj72363 Streptomy

ALIGNMENTS

RESULT 1
AAD04333
ID AAD04333 standard; cDNA; 1782 BP.
XX
AC AAD04333;
XX
DT 04-JUL-2001 (first entry)
XX
DE Human TNF receptor-associated factor (TRAF2) binding protein, IREN cDNA.
XX
KW Human; Tumour Necrosis Factor; TNF; TNF Receptor Associated Factor;
KW TRAF2; TRAF2 binding protein; IREN; IkappaB Regulator; immunosuppressive;
KW nuclear factor-kappaB; NF-kappaB; cytostatic; tumour; AIDS;
KW acquired immune deficiency syndrome; rheumatic disease; apoptosis;
KW autoimmune disease; septic shock; graft-vs-host reaction; inflammation;
KW anorexia; anti-HIV; therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1. .1626
FT /*tag= a
FT /product= "Human TRAF2-binding protein, IREN"
XX
PN WO200116314-A1.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-IL000517.
XX
PR 02-SEP-1999; 99IL-00131719.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Wallach D, Malinin N, Sinha I, Leu S;
XX
DR WPI; 2001-281387/29.
DR P-PSDB; AAE00683.
XX
PT New DNA sequence encoding Tumor Necrosis Factor receptor associated
PT factor (TRAF) binding proteins (IREN) for treatment or prevention of
PT pathological conditions associated with NF-kappaB induction.
XX

PS Claim 1; Fig 3B; 118pp; English.
XX
CC The present cDNA sequence encodes human tumour necrosis factor (TNF)
CC receptor-associated factor (TRAF2) binding protein, IREN. A fragment of
CC this IREN molecule is capable of binding to TRAF2 protein at position 225
CC -501. The invention relates to human tumour necrosis factor (TNF)
CC receptor- associated factor (TRAF2) binding protein designated as IREN
CC (IkappaB REgulator), its isoforms IREN-10B, IREN-E and their
CC corresponding cDNA molecules. IREN is useful for modulating/mediating the
CC activity of transcription factor NF (Nuclear Factor)-kappaB or any other
CC intracellular signalling activity mediated by TRAF2. IREN is useful in
CC the prevention and treatment of a pathological condition associated with
CC NF-kappaB induction (abnormal) e.g. AIDS (acquired immune deficiency
CC syndrome), autoimmune diseases, tumours, rheumatic diseases, anorexia,
CC septic shock and graft-vs-host reactions. IREN also plays an important
CC role in the control of inflammation and other non-apoptotic effects of
CC TNF as well as in the control of apoptosis. The invention also relates to
CC method for screening, identifying and producing a molecule capable of
CC modulating activities mediated by IREN. IREN antibodies are useful for
CC the purification of new proteins from different sources, including cell
CC extracts or transformed cell lines, in addition IREN can be used in
CC diagnostic purposes for identifying disorders related to abnormal
CC functioning of cellular effects mediated directly by TRAF proteins
XX
SQ Sequence 1782 BP; 456 A; 458 C; 504 G; 364 T; 0 U; 0 Other;

Query Match 100.0%; Score 1782; DB 4; Length 1782;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1782; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60

Qy         61 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 120

Qy        121 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 180

Qy        181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 240

Qy        241 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 300

Qy        301 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 360

Qy        361 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 420

Qy        421 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 480

Qy        481 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 540

Qy        541 GATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 600

Qy        601 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 660

Qy        661 ATCACAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATCACAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTG 720

Qy        721 CCTGTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTGTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAA 780

Qy        781 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 840

Qy        841 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCT 900

Qy        901 GTCAATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GTCAATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGC 960

Qy        961 AGCAACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGCAACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTT 1020

Qy       1021 GATGTGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GATGTGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAAC 1080

Qy       1081 TCATCAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAAGCCACTGGAAGGGAAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCATCAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAAGCCACTGGAAGGGAAC 1140

Qy       1141 ACCTGCCTCTCCCAGATGCACAGCTGGGCTCCGCTGAAGGTGCTGCACAATGACTCCGAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACCTGCCTCTCCCAGATGCACAGCTGGGCTCCGCTGAAGGTGCTGCACAATGACTCCGAC 1200

Qy       1201 ATCCTCTTCCCTGTCAGTGGCGTGGGCTCCTACAGCCCAGCAGATGCCCCCCCTCGGAAGC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATCCTCTTCCCTGTCAGTGGCGTGGGCTCCTACAGCCCAGCAGATGCCCCCCTCGGAAGC 1260

Qy       1261 CTGGAGAACGGGACAGGACCAGAGGACCACGTTCTCCCGGATCCTGGACTTCGGTACAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTGGAGAACGGGACAGGACCAGAGGACCACGTTCTCCCGGATCCTGGACTTCGGTACAGT 1320

Qy       1321 GTGGAAGCCAGCTCTCCAGGCCACGGAAGTCCTCTGAGCAGCCTGTTACCTTCTGCCTCA 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTGGAAGCCAGCTCTCCAGGCCACGGAAGTCCTCTGAGCAGCCTGTTACCTTCTGCCTCA 1380

Qy       1381 GTGCCAGAGTCCATGACAATTAGTGAACTGCGCCAGGCCACTGTGGCCATGATGAACAGG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GTGCCAGAGTCCATGACAATTAGTGAACTGCGCCAGGCCACTGTGGCCATGATGAACAGG 1440

Qy       1441 AAGGATGAGCTGGAGGAGGAGAACAGATCACTGCGAAACCTGCTCGACGGTGAGATGGAG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AAGGATGAGCTGGAGGAGGAGAACAGATCACTGCGAAACCTGCTCGACGGTGAGATGGAG 1500

Qy       1501 CACTCAGCCGCGCTCCGGCAAGAGGTGGACACCTTGAAAAGGAAGGTGGCTGAACAGGAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CACTCAGCCGCGCTCCGGCAAGAGGTGGACACCTTGAAAAGGAAGGTGGCTGAACAGGAG 1560

Qy       1561 GAGCGGCAGGGCATGAAGGTCCAGGCGCTGGCCAGCTATCTTTGCTATTTTGTGAGGAGA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GAGCGGCAGGGCATGAAGGTCCAGGCGCTGGCCAGCTATCTTTGCTATTTTGTGAGGAGA 1620

Qy       1621 TTCTAACCCCACGTGAGAACCATGTGGTGGAGAAATGGAGGGAGAGAGAAATCCAACAGT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TTCTAACCCCACGTGAGAACCATGTGGTGGAGAAATGGAGGGAGAGAGAAATCCAACAGT 1680

Qy       1681 TCCTGATAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCCTGAAGCTGTGTTTCCTCTGG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TCCTGATAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCCTGAAGCTGTGTTTCCTCTGG 1740
```

```
Qy       1741 ACTTTTCATGTATGTGAGCCAATAAATTGCTTTCATTCCTTG 1782
             ||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ACTTTTCATGTATGTGAGCCAATAAATTGCTTTCATTCCTTG 1782


RESULT 2
AAV03323
ID   AAV03323 standard; cDNA; 1906 BP.
XX
AC   AAV03323;
XX
DT   15-APR-1998  (first entry)
XX
DE   5' end of clone 9, which encodes a TRAF2 binding protein.
XX
KW   Human tumour necrosis factor receptor-associated factor 2; TRAF2;
KW   TRAF-2 binding protein; NF-kappaB activity; NF-kappaB induction;
KW   intracellular signalling activity; acute hepatitis;
KW   autoimmune-induced cell death; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             94. .1906
FT                   /*tag=  a
FT                   /codon= 127-129, aa:Ser
FT                   /note= "no ATG start or STOP codon given"
XX
PN   WO9737016-A1.
XX
PD   09-OCT-1997.
XX
PF   01-APR-1997;   97WO-IL000117.
XX
PR   02-APR-1996;   96IL-00117800.
PR   26-AUG-1996;   96IL-00119133.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Wallach D,  Malinin N,  Boldin M,  Kovalenko A,  Mett I;
XX
DR   WPI; 1997-503101/46.
DR   P-PSDB; AAW42400.
XX
PT   DNA encoding tumour necrosis factor receptor-associated factor binding
PT   molecule - used for modulation or mediation in cells of the activity of
PT   NF-kB.
XX
PS   Claim 4; Fig 3; 127pp; English.
XX
CC   The present sequence represents the 5' end of clone 9, which encodes a
CC   TRAF2 binding protein. Clone 9 is a partial clone, which lacks most of
CC   its 5' end of the coding DNA sequence. A cDNA library prepared from B-
CC   cells was screened for proteins that associate with TRAF2, and the
CC   present sequence isolated. The clone 9 protein is capable of binding to
CC   at least amino acids 222-501 of TRAF2. The TRAF-2 binding proteins can be
CC   used for modulation or mediation in cells of the activity of NF-kappaB or
CC   any other intracellular signalling activity modulated or mediated by
CC   TRAF2. TRAF-binding proteins are especially used for prevention or
CC   treatment of pathological conditions associated with NF-kB induction,
CC   e.g. acute hepatitis, autoimmune-induced cell death, e.g. death of the
CC   beta Langerhans cells or the pancreas that results in diabetes, the death
CC   of cells in graft rejection, the death of oligodendrocytes in the brain
CC   in multiple sclerosis, and AIDS-inhibited T cell suicide which causes
CC   proliferation of the AIDS virus and hence the AIDS disease. The proteins
CC   are also useful for screening of ligands capable of binding to a protein,
CC   which are useful for modulating cellular activity modulated/mediated by
CC   TRAF2
XX
SQ   Sequence 1906 BP; 482 A; 492 C; 547 G; 378 T; 0 U; 7 Other;

  Query Match             97.8%;  Score 1742;  DB 2;  Length 1906;
  Best Local Similarity   99.6%;  Pred. No. 0;

  Matches 1775;  Conservative    0;  Mismatches    4;  Indels    3;  Gaps    3;

Qy          1 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60
              |||||||||||  | ||||  ||||||||||||||||||||||||||||||||||||||
Db        118 ATGAGCGGATCNCNGAACNATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 177

Qy         61 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 120
              ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        178 GTGAAACAGTGCCAGATCCGCTTTNGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 237

Qy        121 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 297

Qy        181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 357

Qy        241 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 417

Qy        301 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 477

Qy        361 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 537

Qy        421 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 597

Qy        481 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 657

Qy        541 GATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 GATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 717

Qy        601 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 777

Qy        661 ATCACAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTG 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        778 ATCACAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGA-CCCTTG 836

Qy        721 CCTGTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 CCTGTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAA 896

Qy        781 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 956

Qy        841 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCT 1016

Qy        901 GTCAATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db       1017 GTCAATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAA-TCAGAGC 1075

Qy        961 AGCAACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1076 AGCAACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTT 1135

Qy       1021 GATGTGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1136 GATGTGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAAC 1195
```

```
Qy       1081 TCATCAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAAGCCACTGGAAGGGAAC 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1196 TCATCAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAAGCCACTGGAAGGGAAC 1255

Qy       1141 ACCTGCCTCTCCCAGATGCACAGCTGGGCTCCGCTGAAGGTGCTGCACAATGACTCCGAC 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1256 ACCTGCCTCTCCCAGATGCACAGCTGGGCTCCGCTGAAGGTGCTGCACAATGACTCCGAC 1315

Qy       1201 ATCCTCTTCCCTGTCAGTGGCGTGGGCTCCTACAGCCCAGCAGATGCCCCCCTCGGAAGC 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1316 ATCCTCTTCCCTGTCAGTGGCGTGGGCTCCTACAGCCCAGCAGATGCCCCCCTCGGAAGC 1375

Qy       1261 CTGGAGAACGGGACAGGACCAGAGGACCACGTTCTCCCGGATCCTGGACTTCGGTACAGT 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1376 CTGGAGAACGGGACAGGACCAGAGGACCACGTTCTCCCGGATCCTGGACTTCGGTACAGT 1435

Qy       1321 GTGGAAGCCAGCTCTCCAGGCCACGGAAGTCCTCTGAGCAGCCTGTTACCTTCTGCCTCA 1380
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       1436 GTGGAAGCCAGCTCTCCAGGCCACGGAAGTCCTCTGAGCAGCCTGTTA-CTTCTGCCTCA 1494

Qy       1381 GTGCCAGAGTCCATGACAATTAGTGAACTGCGCCAGGCCACTGTGGCCATGATGAACAGG 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1495 GTGCCAGAGTCCATGACAATTAGTGAACTGCGCCAGGCCACTGTGGCCATGATGAACAGG 1554

Qy       1441 AAGGATGAGCTGGAGGAGGAGAACAGATCACTGCGAAACCTGCTCGACGGTGAGATGGAG 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1555 AAGGATGAGCTGGAGGAGGAGAACAGATCACTGCGAAACCTGCTCGACGGTGAGATGGAG 1614

Qy       1501 CACTCAGCCGCGCTCCGGCAAGAGGTGGACACCTTGAAAAGGAAGGTGGCTGAACAGGAG 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1615 CACTCAGCCGCGCTCCGGCAAGAGGTGGACACCTTGAAAAGGAAGGTGGCTGAACAGGAG 1674

Qy       1561 GAGCGGCAGGGCATGAAGGTCCAGGCGCTGGCCAGCTATCTTTGCTATTTTGTGAGGAGA 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1675 GAGCGGCAGGGCATGAAGGTCCAGGCGCTGGCCAGCTATCTTTGCTATTTTGTGAGGAGA 1734

Qy       1621 TTCTAACCCCACGTGAGAACCATGTGGTGGAGAAATGGAGGGAGAGAGAAATCCAACAGT 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1735 TTCTAACCCCACGTGAGAACCATGTGGTGGAGAAATGGAGGGAGAGAGAAATCCAACAGT 1794

Qy       1681 TCCTGATAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCCTGAAGCTGTGTTTCCTCTGG 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1795 TCCTGATAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCCTGAAGCTGTGTTTCCTCTGG 1854

Qy       1741 ACTTTTCATGTATGTGAGCCAATAAATTGCTTTCATTCCTTG 1782
             |||||||||||||||||||||||||||||||||||||||||||
Db       1855 ACTTTTCATGTATGTGAGCCAATAAATTGCTTTCATTCCTTG 1896
```

```
RESULT 3
AAD04335
ID   AAD04335 standard; cDNA; 2873 BP.
XX
AC   AAD04335;
XX
DT   04-JUL-2001  (first entry)
XX
DE   Human TNF receptor-associated factor binding protein, IREN-E cDNA.
XX
KW   Human; Tumour Necrosis Factor; TNF; TNF Receptor Associated Factor;
KW   TRAF2; TRAF2 binding protein; IREN; IkappaB Regulator; IREN-E;
KW   immunosuppressive; nuclear factor-kappaB; NF-kappaB; cytostatic; tumour;
KW   AIDS; acquired immune deficiency syndrome; rheumatic disease; apoptosis;
KW   autoimmune disease; septic shock; graft-vs-host reaction; inflammation;
KW   anorexia; anti-HIV; therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .2355
FT                   /*tag= a
FT                   /product= "Human TRAF2-binding protein isoform, IREN-E"
XX
PN   WO200116314-A1.
XX
PD   08-MAR-2001.
XX
PF   31-AUG-2000; 2000WO-IL000517.
XX
PR   02-SEP-1999;   99IL-00131719.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Wallach D,  Malinin N,  Sinha I,  Leu S;
XX
DR   WPI; 2001-281387/29.
DR   P-PSDB; AAE00685.
XX
PT   New DNA sequence encoding Tumor Necrosis Factor receptor associated
PT   factor (TRAF) binding proteins (IREN) for treatment or prevention of
PT   pathological conditions associated with NF-kappaB induction.
XX
PS   Claim 1; Fig 5; 118pp; English.
XX
CC   The present cDNA sequence encodes human tumour necrosis factor (TNF)
CC   receptor-associated factor (TRAF2) binding protein isoform, IREN-E. A
CC   fragment of this IREN-E molecule is capable of binding to TRAF2 protein
CC   at position 225-501. The invention relates to human tumour necrosis
CC   factor (TNF) receptor- associated factor (TRAF2) binding protein
CC   designated as IREN (IkappaB REgulator), its isoforms IREN-10B, IREN-E and
CC   their corresponding cDNA molecules. IREN is useful for
CC   modulating/mediating the activity of transcription factor NF (Nuclear
CC   Factor)-kappaB or any other intracellular signalling activity mediated by
CC   TRAF2. IREN is useful in the prevention and treatment of a pathological
CC   condition associated with NF-kappaB induction (abnormal) e.g. AIDS
CC   (acquired immune deficiency syndrome), autoimmune diseases, tumours,
CC   rheumatic diseases, anorexia, septic shock and graft-vs-host reactions.
CC   IREN also plays an important role in the control of inflammation and
CC   other non-apoptotic effects of TNF as well as in the control of
CC   apoptosis. The invention also relates to method for screening,
CC   identifying and producing a molecule capable of modulating activities
CC   mediated by IREN. IREN antibodies are useful for the purification of new
CC   proteins from different sources, including cell extracts or transformed
CC   cell lines, in addition IREN can be used in diagnostic purposes for
CC   identifying disorders related to abnormal functioning of cellular effects
CC   mediated directly by TRAF proteins
XX
SQ   Sequence 2873 BP; 756 A; 736 C; 806 G; 575 T; 0 U; 0 Other;
```

```
  Query Match           89.5%;  Score 1595;  DB 4;  Length 2873;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1595;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60

Qy         61 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 120

Qy        121 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 180

Qy        181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 240

Qy        241 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 300

Qy        301 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        301 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 360

Qy        361 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 420

Qy        421 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 480

Qy        481 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 540

Qy        541 GATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 600

Qy        601 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 660

Qy        661 ATCACAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATCACAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTG 720

Qy        721 CCTGTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTGTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAA 780

Qy        781 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 840

Qy        841 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCT 900

Qy        901 GTCAATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GTCAATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGC 960

Qy        961 AGCAACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGCAACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTT 1020

Qy       1021 GATGTGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GATGTGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAAC 1080

Qy       1081 TCATCAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAAGCCACTGGAAGGGAAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCATCAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAAGCCACTGGAAGGGAAC 1140

Qy       1141 ACCTGCCTCTCCCAGATGCACAGCTGGGCTCCGCTGAAGGTGCTGCACAATGACTCCGAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACCTGCCTCTCCCAGATGCACAGCTGGGCTCCGCTGAAGGTGCTGCACAATGACTCCGAC 1200

Qy       1201 ATCCTCTTCCCTGTCAGTGGCGTGGGCTCCTACAGCCCAGCAGATGCCCCCCTCGGAAGC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATCCTCTTCCCTGTCAGTGGCGTGGGCTCCTACAGCCCAGCAGATGCCCCCCTCGGAAGC 1260

Qy       1261 CTGGAGAACGGGACAGGACCAGAGGACCACGTTCTCCCGGATCCTGGACTTCGGTACAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTGGAGAACGGGACAGGACCAGAGGACCACGTTCTCCCGGATCCTGGACTTCGGTACAGT 1320

Qy       1321 GTGGAAGCCAGCTCTCCAGGCCACGGAAGTCCTCTGAGCAGCCTGTTACCTTCTGCCTCA 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTGGAAGCCAGCTCTCCAGGCCACGGAAGTCCTCTGAGCAGCCTGTTACCTTCTGCCTCA 1380

Qy       1381 GTGCCAGAGTCCATGACAATTAGTGAACTGCGCCAGGCCACTGTGGCCATGATGAACAGG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GTGCCAGAGTCCATGACAATTAGTGAACTGCGCCAGGCCACTGTGGCCATGATGAACAGG 1440

Qy       1441 AAGGATGAGCTGGAGGAGGAGAACAGATCACTGCGAAACCTGCTCGACGGTGAGATGGAG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AAGGATGAGCTGGAGGAGGAGAACAGATCACTGCGAAACCTGCTCGACGGTGAGATGGAG 1500

Qy       1501 CACTCAGCCGCGCTCCGGCAAGAGGTGGACACCTTGAAAAGGAAGGTGGCTGAACAGGAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CACTCAGCCGCGCTCCGGCAAGAGGTGGACACCTTGAAAAGGAAGGTGGCTGAACAGGAG 1560

Qy       1561 GAGCGGCAGGGCATGAAGGTCCAGGCGCTGGCCAG 1595
              ||||||||||||||||||||||||||||||||||||
Db       1561 GAGCGGCAGGGCATGAAGGTCCAGGCGCTGGCCAG 1595
```

RESULT 4
AAD04334

```
ID   AAD04334 standard; cDNA; 3139 BP.
XX
AC   AAD04334;
XX
DT   04-JUL-2001  (first entry)
XX
DE   Human TNF receptor-associated factor binding protein, IREN-10B cDNA.
XX
KW   Human; Tumour Necrosis Factor; TNF; TNF Receptor Associated Factor;
KW   TRAF2; TRAF2 binding protein; IkappaB Regulator; IREN-10B;
KW   immunosuppressive; nuclear factor-kappaB; NF-kappaB; cytostatic; tumour;
KW   AIDS; acquired immune deficiency syndrome; rheumatic disease; apoptosis;
KW   autoimmune disease; septic shock; graft-vs-host reaction; inflammation;
KW   anorexia; anti-HIV; therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .2442
FT                   /*tag=  a
FT                   /product= "Human TRAF2-binding protein isoform, IREN-10B"
FT
XX
PN   WO200116314-A1.
XX
PD   08-MAR-2001.
XX
PF   31-AUG-2000; 2000WO-IL000517.
XX
PR   02-SEP-1999;   99IL-00131719.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Wallach D,  Malinin N,  Sinha I,  Leu S;
XX
DR   WPI; 2001-281387/29.
DR   P-PSDB; AAE00684.
XX
PT   New DNA sequence encoding Tumor Necrosis Factor receptor associated
PT   factor (TRAF) binding proteins (IREN) for treatment or prevention of
PT   pathological conditions associated with NF-kappaB induction.
XX
PS   Claim 1; Fig 4; 118pp; English.
XX
CC   The present cDNA sequence encodes human tumour necrosis factor (TNF)
CC   receptor-associated factor (TRAF2) binding protein isoform, IREN-10B. A
CC   fragment of this IREN-10B molecule is capable of binding to TRAF2 protein
CC   at position 225-501. The invention relates to human tumour necrosis
CC   factor (TNF) receptor- associated factor (TRAF2) binding protein
CC   designated as IREN (IkappaB REgulator), its isoforms IREN-10B, IREN-E and
CC   their corresponding cDNA molecules. IREN is useful for
CC   modulating/mediating the activity of transcription factor NF (Nuclear
CC   Factor)-kappaB or any other intracellular signalling activity mediated by
CC   TRAF2. IREN is useful in the prevention and treatment of a pathological
CC   condition associated with NF-kappaB induction (abnormal) e.g. AIDS
CC   (acquired immune deficiency syndrome), autoimmune diseases, tumours,
CC   rheumatic diseases, anorexia, septic shock and graft-vs-host reactions.
CC   IREN also plays an important role in the control of inflammation and
```

```
CC   other non-apoptotic effects of TNF as well as in the control of
CC   apoptosis. The invention also relates to method for screening,
CC   identifying and producing a molecule capable of modulating activities
CC   mediated by IREN. IREN antibodies are useful for the purification of new
CC   proteins from different sources, including cell extracts or transformed
CC   cell lines, in addition IREN can be used in diagnostic purposes for
CC   identifying disorders related to abnormal functioning of cellular effects
CC   mediated directly by TRAF proteins
XX
SQ   Sequence 3139 BP; 797 A; 851 C; 875 G; 616 T; 0 U; 0 Other;

  Query Match             89.5%;  Score 1595;  DB 4;  Length 3139;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1595;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60

Qy         61 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 120

Qy        121 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 180

Qy        181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 240

Qy        241 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 300

Qy        301 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 360

Qy        361 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 420

Qy        421 AGGCTGAGCACTTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGCTGAGCACTTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 480

Qy        481 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 540

Qy        541 GATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 600

Qy        601 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 660

Qy        661 ATCACAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATCACAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTG 720

Qy        721 CCTGTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTGTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAA 780

Qy        781 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 840

Qy        841 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCT 900

Qy        901 GTCAATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GTCAATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGC 960

Qy        961 AGCAACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGCAACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTT 1020

Qy       1021 GATGTGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GATGTGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAAC 1080

Qy       1081 TCATCAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAAGCCACTGGAAGGGAAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCATCAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAAGCCACTGGAAGGGAAC 1140

Qy       1141 ACCTGCCTCTCCCAGATGCACAGCTGGGCTCCGCTGAAGGTGCTGCACAATGACTCCGAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACCTGCCTCTCCCAGATGCACAGCTGGGCTCCGCTGAAGGTGCTGCACAATGACTCCGAC 1200

Qy       1201 ATCCTCTTCCCTGTCAGTGGCGTGGGCTCCTACAGCCCAGCAGATGCCCCCCTCGGAAGC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATCCTCTTCCCTGTCAGTGGCGTGGGCTCCTACAGCCCAGCAGATGCCCCCCTCGGAAGC 1260

Qy       1261 CTGGAGAACGGGACAGGACCAGAGGACCACGTTCTCCCGGATCCTGGACTTCGGTACAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTGGAGAACGGGACAGGACCAGAGGACCACGTTCTCCCGGATCCTGGACTTCGGTACAGT 1320

Qy       1321 GTGGAAGCCAGCTCTCCAGGCCACGGAAGTCCTCTGAGCAGCCTGTTACCTTCTGCCTCA 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTGGAAGCCAGCTCTCCAGGCCACGGAAGTCCTCTGAGCAGCCTGTTACCTTCTGCCTCA 1380

Qy       1381 GTGCCAGAGTCCATGACAATTAGTGAACTGCGCCAGGCCACTGTGGCCATGATGAACAGG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GTGCCAGAGTCCATGACAATTAGTGAACTGCGCCAGGCCACTGTGGCCATGATGAACAGG 1440

Qy       1441 AAGGATGAGCTGGAGGAGGAGAACAGATCACTGCGAAACCTGCTCGACGGTGAGATGGAG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AAGGATGAGCTGGAGGAGGAGAACAGATCACTGCGAAACCTGCTCGACGGTGAGATGGAG 1500

Qy       1501 CACTCAGCCGCGCTCCGGCAAGAGGTGGACACCTTGAAAAGGAAGGTGGCTGAACAGGAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CACTCAGCCGCGCTCCGGCAAGAGGTGGACACCTTGAAAAGGAAGGTGGCTGAACAGGAG 1560

Qy       1561 GAGCGGCAGGGCATGAAGGTCCAGGCGCTGGCCAG 1595
              ||||||||||||||||||||||||||||||||||||
Db       1561 GAGCGGCAGGGCATGAAGGTCCAGGCGCTGGCCAG 1595


RESULT 5
AAH15665
ID   AAH15665 standard; cDNA; 2248 BP.
XX
AC   AAH15665;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA sequence SEQ ID NO:14016.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-00116126.
XX
PR   29-JUL-1999;   99JP-00248036.
PR   27-AUG-1999;   99JP-00300253.
```

```
PR   11-JAN-2000; 2000JP-00118776.
PR   02-MAY-2000; 2000JP-00183767.
PR   09-JUN-2000; 2000JP-00241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI   Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT   length cDNAs defined in the specification, and for the detection and/or
PT   diagnosis of the abnormality of the proteins encoded by the full-length
PT   cDNAs.
XX
PS   Claim 8; SEQ ID NO 14016; 2537pp + Sequence Listing; English.
XX
CC   The present invention describes primer sets for synthesising 5602 full-
CC   length cDNAs defined in the specification. Where a primer set comprises:
CC   (a) an oligo-dT primer and an oligonucleotide complementary to the
CC   complementary strand of a polynucleotide which comprises one of the 5602
CC   nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in the
CC   specification. The primer sets can be used in antisense therapy and in
CC   gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC   represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC   oligonucleotides, all of which are used in the exemplification of the
CC   present invention
XX
SQ   Sequence 2248 BP; 548 A; 548 C; 618 G; 534 T; 0 U; 0 Other;

  Query Match             63.4%;  Score 1129.8;  DB 4;  Length 2248;
  Best Local Similarity   95.3%;  Pred. No. 1.1e-297;
  Matches 1164;  Conservative    0;  Mismatches   57;  Indels    0;  Gaps    0;

Qy          1 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 122

Qy         61 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 182

Qy        121 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 242

Qy        181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 302

Qy        241 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 362

Qy        301 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 422

Qy        361 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 482

Qy        421 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 542

Qy        481 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 602

Qy        541 GATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 600
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 GATTTGGACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 662

Qy        601 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 722

Qy        661 ATCACAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 ATCACAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTG 782

Qy        721 CCTGTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 CCTGTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAG 842

Qy        781 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        843 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 902

Qy        841 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        903 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCT 962

Qy        901 GTCAATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 GTCAATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGC 1022

Qy        961 AGCAACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 AGCAACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTT 1082

Qy       1021 GATGTGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAAC 1080
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1083 GATGTGAAAGGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAAC 1142

Qy       1081 TCATCAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAAGCCACTGGAAGGGAAC 1140
              |||||||||||||||||||||||||||||||||||||||||||||  | |    | | |
Db       1143 TCATCAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAAGTAAGTTTGTGTGTAA 1202

Qy       1141 ACCTGCCTCTCCCAGATGCACAGCTGGGCTCCGCTGAAGGTGCTGCACAATGACTCCGAC 1200
                  |  |||| | |   | |||   ||      |    ||| |      | ||| | ||
Db       1203 GGTGGAGTCTCACCGTCCCCCAGGCTGGAGTGCCGTGGCGTGATCTCAGCTCACTGCAAC 1262

Qy       1201 ATCCTCTTCCCTGTCAGTGGC 1221
                |||  | |||| |    |||
Db       1263 CTCCACCTCCCGGGTTCAGGC 1283


RESULT 6
AAH18235
ID   AAH18235 standard; cDNA; 6045 BP.
XX
AC   AAH18235;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA sequence SEQ ID NO:18174.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
```

PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAn.
XX
PS Claim 8; SEQ ID NO 18174; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 6045 BP; 1561 A; 1493 C; 1427 G; 1564 T; 0 U; 0 Other;

Query Match 45.8%; Score 816.8; DB 4; Length 6045;
Best Local Similarity 97.0%; Pred. No. 6e-212;
Matches 854; Conservative 0; Mismatches 22; Indels 4; Gaps 2;

```
Qy        246 AGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGC-AGCGCT 304
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        932 AAAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAAGCGCT 991

Qy        305 TCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGCTGTG 364
              |||||||||||||||||||||||||||||||||||||| ||| ||||||||||||||||
Db        992 TCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGCCGCACCTGGCTGCGCTGTG 1051

Qy        365 CCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGCAGGC 424
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1052 CCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGCAGGC 1111

Qy        425 TGAGCACTTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTTC 484
              | || ||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       1112 TCAGTACTTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAGAGGTCCAGTATGCTTC 1171

Qy        485 CTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAGGATT 544
              |||||||||||||||||| |||||||||||||||||||||||||| ||||||||||||
Db       1172 CTACCATGGCAGCAGGTCCGAACTCCATACTCTTTGCGATTAACATTGACAACAAGGATT 1231

Qy        545 TGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGA 604
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1232 TGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGA 1291

Qy        605 ACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCA 664
              | ||||  ||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       1292 ATGTGA---CCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCGTGTTCAGGGAGATCA 1348

Qy        665 CAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCTG 724
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||| |
Db       1349 CAGCCTCCTCTGCCATCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCCG 1408

Qy        725 TCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGA 784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1409 TCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGC 1468

Qy        785 AAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGT 844
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||| |||
Db       1469 AAGTGACCAACATTATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACATGT 1528

Qy        845 TTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCA 904
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       1529 TTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCACTCCTCTGTCA 1588

Qy        905 ATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGCAGCA 964
              |||||||||||||||||||||||||||||||||  |||||||||||||||||||||||||
Db       1589 ATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCGAAACTCCAATGGAAGTCAGAGCAGCA 1648

Qy        965 ACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTTGATG 1024
              |||| |||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       1649 ACTCGTGGAAAATTGATTCCCTGTCTTTGAACAGGGAGTTTGGGTACCAGAAGCTTGATG 1708

Qy       1025 TGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCAT 1084
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1709 TGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCAT 1768

Qy       1085 CAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAA 1124
              ||||| |||||||||||||||||| ||||||||||||||
Db       1769 CAGGACGGAAGCACAGGGGCCACTCAGAGTCGCCCGAGAA 1808
```

RESULT 7
AAH04973
ID AAH04973 standard; cDNA; 730 BP.
XX
AC AAH04973;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:1808.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

```
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT   length cDNAs defined in the specification, and for the detection and/or
PT   diagnosis of the abnormality of the proteins encoded by the full-length
PT   cDNAs.
XX
PS   Claim 1; SEQ ID NO 1808; 2537pp + Sequence Listing; English.
XX
CC   The present invention describes primer sets for synthesising 5602 full-
CC   length cDNAs defined in the specification. Where a primer set comprises:
CC   (a) an oligo-dT primer and an oligonucleotide complementary to the
CC   complementary strand of a polynucleotide which comprises one of the 5602
CC   nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in the
CC   specification. The primer sets can be used in antisense therapy and in
CC   gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC   represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC   oligonucleotides, all of which are used in the exemplification of the
CC   present invention
XX
SQ   Sequence 730 BP; 173 A; 200 C; 217 G; 137 T; 0 U; 3 Other;

  Query Match           35.9%;  Score 639.8;  DB 4;  Length 730;
  Best Local Similarity  99.0%;  Pred. No. 5e-164;
  Matches  663;  Conservative    0;  Mismatches    5;  Indels    2;  Gaps    2;

Qy          1 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 122

Qy         61 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 182

Qy        121 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 242

Qy        181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 302

Qy        241 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 362

Qy        301 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 422

Qy        361 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 482

Qy        421 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 542

Qy        481 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 602

Qy        541 GATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 GATTTGGACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 662

Qy        601 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 660
              || ||||||| || ||||||||||||||||||||||||||||||| || |||||||| ||||
Db        663 CANAACGTGA-CTNCTTGCTGAAGGAGTCCACGCAAGGAGTGANCA-CCTGTTCAAGGAG 720

Qy        661 ATCACAGCCT 670
              ||||||||||
Db        721 ATCACAGCCT 730


RESULT 8
AAK67282
ID   AAK67282 standard; DNA; 33147 BP.
XX
AC   AAK67282;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:22094.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001354.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
```

PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS Disclosure; SEQ ID NO 22094; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 33147 BP; 8370 A; 7682 C; 8382 G; 8713 T; 0 U; 0 Other;

Query Match 20.4%; Score 364.2; DB 4; Length 33147;
Best Local Similarity 96.6%; Pred. No. 6.8e-88;
Matches 372; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 746 CTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGAAAGTGACCAACATAATCTCAT 805
Db 8894 CAGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGCAAGTGACCAACATTAACTCAT 8953

Qy 806 TTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGTTTAAAAAGACACCTGGGGCAG 865
Db 8954 TTGATGATGAGGAAGATGAGCAGAACTCTGGGGACATGTTTAAAAAGACACCTGGGGCAG 9013

Qy 866 GGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCAATATCATGTCCGCCTTTGAAA 925
Db 9014 GGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCAATATCATGTCCGCCTTTGAAA 9073

```
Qy       926 GCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGCAGCAACTCATGGAAAATTGATTCCC 985
             ||||||||||||| ||||||||||||||||||||||||||||| ||||||||||||||||
Db      9074 GCCCCTTCGGGCCAAACTCCAATGGAAGTCAGAGCAGCAACTCGTGGAAAATTGATTCCC 9133

Qy       986 TGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTTGATGTGAAAAGCATCGATGATGAAG 1045
             |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db      9134 TGTCTTTGAACAGGGAGTTTGGGTACCAGAAGCTTGATGTGAAAAGCATCGATGATGAAG 9193

Qy      1046 ATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCATCAGGAAGGAAGCACAGGGGCC 1105
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db      9194 ATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCATCAGGACGGAAGCACAGGGGCC 9253

Qy      1106 ACTCGGAGTCGCCCGAGAAGCCACT 1130
             |||| ||||||||||||||||  | |
Db      9254 ACTCAGAGTCGCCCGAGAAGTAAGT 9278
```

RESULT 9
ABS72233
ID ABS72233 standard; cDNA; 280 BP.
XX
AC ABS72233;
XX
DT 03-DEC-2002 (first entry)
XX
DE Human gene trapped sequence (GTS) #193.
XX
KW Human; gene trapped sequence; GTS; gene; ss; cancer; autoimmune disease;
KW lupus; scleroderma; Crohn's disease; multiple sclerosis; immune disorder;
KW inflammatory bowel disease; schizophrenia; psychosis; osteoarthritis;
KW inflammatory disorder; diabetes; skin disorder; acne; eczema; asthma;
KW rheumatoid arthritis; hypertension; atherosclerosis; Alzheimer's disease;
KW cardiovascular disease; Parkinson's disease; osteoporosis; infertility;
KW viral infection; parasitic infection; fungal infection;
KW bacterial infection; forensic analysis; cellular differentiation.
XX
OS Homo sapiens.
XX
PN US2002095031-A1.
XX
PD 18-JUL-2002.
XX
PF 03-MAY-2000; 2000US-00563817.
XX
PR 04-MAY-1999; 99US-0132343P.
XX
PA (NEHL/) NEHLS M C.
PA (ZAMB/) ZAMBROWICZ B.
PA (SAND/) SANDS A T.
XX
PI Nehls MC, Zambrowicz B, Sands AT;
XX
DR WPI; 2002-656030/70.
XX
PT New isolated or purified human gene trapped sequences, useful for gene
PT discovery, as markers for gene expression analysis, identifying and
PT mapping the coding regions of human genome, or determining the genetic
PT basis of human disease.
XX
PS Claim 1; SEQ ID NO 201; 36pp; English.
XX
CC The invention relates to isolated or purified polynucleotides that
CC correspond to human gene trapped sequences (GTSs). The human GTSs are
CC useful for gene discovery and as markers for gene expression analysis,
CC for identifying and mapping the coding regions of the mammalian,
CC particularly human, genome, for forensic analysis, and for determining
CC the genetic basis of human disease. The peptides and proteins encoded by
CC the polynucleotides are useful for generating antibodies, as reagents in
CC diagnostic assays and in identifying other cellular gene products
CC involved in the regulation of development and cellular differentiation of
CC various cell types. The peptides are also useful as reagents in assays
CC for screening of compounds used in treating disorders affecting
CC development and cell differentiation. The GTSs are also useful in
CC treating or ameliorating diseases associated with the expression of
CC mutant or normal variants of the GTSs, e.g. cancer, autoimmune diseases,
CC lupus, scleroderma, Crohn's disease, multiple sclerosis, inflammatory
CC bowel disease, immune disorders, schizophrenia, psychosis, inflammatory
CC disorders, diabetes, skin disorders such as acne or eczema,
CC osteoarthritis, rheumatoid arthritis, hypertension, atherosclerosis,
CC cardiovascular diseases, Alzheimer's disease, Parkinson's disease,
CC osteoporosis, asthma, infertility, and viral, parasitic, fungal or
CC bacterial infections. This sequence represents a human GTS of the
CC invention
XX
SQ Sequence 280 BP; 71 A; 66 C; 91 G; 51 T; 0 U; 1 Other;

```
  Query Match            15.5%;  Score 275.8;  DB 6;  Length 280;
  Best Local Similarity  98.9%;  Pred. No. 9.2e-65;
  Matches  277;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1349 GTCCTCTGAGCAGCCTGTTACCTTCTGCCTCAGTGCCAGAGTCCATGACAATTAGTGAAC 1408
             | ||| |||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         1 GGCCTATGAGCAGCCTGNTACCTTCTGCCTCAGTGCCAGAGTCCATGACAATTAGTGAAC 60

Qy      1409 TGCGCCAGGCCACTGTGGCCATGATGAACAGGAAGGATGAGCTGGAGGAGGAGAACAGAT 1468
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 TGCGCCAGGCCACTGTGGCCATGATGAACAGGAAGGATGAGCTGGAGGAGGAGAACAGAT 120

Qy      1469 CACTGCGAAACCTGCTCGACGGTGAGATGGAGCACTCAGCCGCGCTCCGGCAAGAGGTGG 1528
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 CACTGCGAAACCTGCTCGACGGTGAGATGGAGCACTCAGCCGCGCTCCGGCAAGAGGTGG 180

Qy      1529 ACACCTTGAAAAGGAAGGTGGCTGAACAGGAGGAGCGGCAGGGCATGAAGGTCCAGGCGC 1588
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 ACACCTTGAAAAGGAAGGTGGCTGAACAGGAGGAGCGGCAGGGCATGAAGGTCCAGGCGC 240

Qy      1589 TGGCCAGCTATCTTTGCTATTTTGTGAGGAGATTCTAACC 1628
             ||||||||||||||||||||||||||||||||||||||||
Db       241 TGGCCAGCTATCTTTGCTATTTTGTGAGGAGATTCTAACC 280
```

RESULT 10
AAS64410
ID AAS64410 standard; cDNA; 454 BP.
XX
AC AAS64410;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #214.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG00223.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess

```
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 214; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 454 BP; 107 A; 110 C; 151 G; 86 T; 0 U; 0 Other;

  Query Match             13.9%;  Score 247;  DB 5;  Length 454;
  Best Local Similarity   100.0%;  Pred. No. 8.7e-57;
  Matches  247;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 127

Qy         61 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 187

Qy        121 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 247

Qy        181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 307

Qy        241 GAAACAG 247
              |||||||
Db        308 GAAACAG 314


RESULT 11
ACH75347/c
ID   ACH75347 standard; DNA; 542 BP.
XX
AC   ACH75347;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human genome derived single exon probe #8542.
XX
KW   Human; probe; ss; gene expression; single exon probe; microarray;
KW   alternative splicing event; genomic alteration.
XX
OS   Homo sapiens.
XX
PN   US2003194704-A1.
XX
PD   16-OCT-2003.
XX
PF   03-APR-2002; 2002US-00029386.
XX
PR   03-APR-2002; 2002US-00029386.
XX
PA   (PENN/) PENN S G.
PA   (RANK/) RANK D R.
PA   (HANZ/) HANZEL D K.
XX
PI   Penn SG,  Rank DR,  Hanzel DK;
XX
DR   WPI; 2004-119264/12.
XX
PT   New human genome-derived single exon nucleic acid probes useful for human
PT   gene expression analysis, for identifying or characterizing alternative
PT   splicing events, for assessing genomic alterations or as tools for
PT   surveying tissues.
XX
PS   Claim 15; SEQ ID NO 8542; 80pp; English.
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
CC   storage medium which contains a database having a plurality of records
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 542 BP; 135 A; 143 C; 129 G; 135 T; 0 U; 0 Other;

  Query Match             12.8%;  Score 227.8;  DB 12;  Length 542;
  Best Local Similarity   96.1%;  Pred. No. 1.7e-51;
  Matches  245;  Conservative    0;  Mismatches    7;  Indels    3;  Gaps    1;

Qy        497 CAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAGGATTTGAACGGGCAGA 556
              |||||| |||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        495 CAGGTCCGAACTCCATACTCTTTGCGATTAACATTGACAACAAGGATTTGAACGGGCAGA 436

Qy        557 GTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAACGTGACCTCCT 616
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||   |||
Db        435 GTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAATGTGA---CCT 379

Qy        617 TGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCACAGCCTCCTCTG 676
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        378 TGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCGTGTTCAGGGAGATCACAGCCTCCTCTG 319

Qy        677 CCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCTGTCGTGTCCAGGA 736
```

```
                || |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        318 CCATCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCCGTCGTGTCCAGGA 259

Qy        737 ATGTCAGTGCTGATG 751
              ||||||||||||| ||
Db        258 ATGTCAGTGCTGGTG 244


RESULT 12
ACH89047/c
ID   ACH89047 standard; DNA; 284 BP.
XX
AC   ACH89047;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human genome derived single exon probe #22242.
XX
KW   Human; probe; ss; gene expression; single exon probe; microarray;
KW   alternative splicing event; genomic alteration.
XX
OS   Homo sapiens.
XX
PN   US2003194704-A1.
XX
PD   16-OCT-2003.
XX
PF   03-APR-2002; 2002US-00029386.
XX
PR   03-APR-2002; 2002US-00029386.
XX
PA   (PENN/) PENN S G.
PA   (RANK/) RANK D R.
PA   (HANZ/) HANZEL D K.
XX
PI   Penn SG,  Rank DR,  Hanzel DK;
XX
DR   WPI; 2004-119264/12.
XX
PT   New human genome-derived single exon nucleic acid probes useful for human
PT   gene expression analysis, for identifying or characterizing alternative
PT   splicing events, for assessing genomic alterations or as tools for
PT   surveying tissues.
XX
PS   Claim 1; SEQ ID NO 22242; 80pp; English.
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
CC   storage medium which contains a database having a plurality of records
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 284 BP; 60 A; 78 C; 74 G; 72 T; 0 U; 0 Other;

  Query Match             12.6%;  Score 224.8;  DB 12;  Length 284;
  Best Local Similarity   96.0%;  Pred. No. 8e-51;
  Matches  242;  Conservative    0;  Mismatches    7;  Indels    3;  Gaps    1;

Qy        500 GTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAGGATTTGAACGGGCAGAGTA 559
              ||| |||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        284 GTCCGAACTCCATACTCTTTGCGATTAACATTGACAACAAGGATTTGAACGGGCAGAGTA 225

Qy        560 AGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAACGTGACCTCCTTGC 619
              |||||||||||||||||||||||||||||||||||||||||||| ||||   ||||||
Db        224 AGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAATGTGA---CCTTGC 168

Qy        620 TGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCACAGCCTCCTCTGCCG 679
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        167 TGAAGGAGTCCACGCAAGGAGTGAGCAGCGTGTTCAGGGAGATCACAGCCTCCTCTGCCA 108

Qy        680 TCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCTGTCGTGTCCAGGAATG 739
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        107 TCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCCGTCGTGTCCAGGAATG 48

Qy        740 TCAGTGCTGATG 751
              |||||||| ||
Db         47 TCAGTGCTGGTG 36


RESULT 13
AAS92200
ID   AAS92200 standard; cDNA; 463 BP.
XX
AC   AAS92200;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #28004.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG28013.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 28004; 103pp; English.
```

```
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 463 BP; 124 A; 133 C; 122 G; 84 T; 0 U; 0 Other;

  Query Match           11.6%;  Score 207.6;  DB 5;  Length 463;
  Best Local Similarity  93.9%;  Pred. No. 5.1e-46;
  Matches  216;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1219 GGCGTGGGCTCCTACAGCCCAGCAGATGCCCCCCTCGGAAGCCTGGAGAACGGGACAGGA 1278
                | | |  || |   | |  ||  ||||||||||||||||||||||||||||||||||||
Db        234 GTCCTCAGCCCTGGCTGAACAAAAGATGCCCCCCTCGGAAGCCTGGAGAACGGGACAGGA 293

Qy       1279 CCAGAGGACCACGTTCTCCCGGATCCTGGACTTCGGTACAGTGTGGAAGCCAGCTCTCCA 1338
                ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        294 CCAGAGGACCACGTTCTCCCAGATCCTGGACTTCGGTACAGTGTGGAAGCCAGCTCTCCA 353

Qy       1339 GGCCACGGAAGTCCTCTGAGCAGCCTGTTACCTTCTGCCTCAGTGCCAGAGTCCATGACA 1398
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        354 GGCCACGGAAGTCCTCTGAGCAGCCTGTTACCTTCTGCCTCAGTGCCAGAGTCCATGACA 413

Qy       1399 ATTAGTGAACTGCGCCAGGCCACTGTGGCCATGATGAACAGGAAGGATGA 1448
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 ATTAGTGAACTGCGCCAGGCCACTGTGGCCATGATGAACAGGAAGGATGA 463


RESULT 14
AAS73549/c
ID   AAS73549 standard; cDNA; 474 BP.
XX
AC   AAS73549;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #9353.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG09362.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 9353; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 474 BP; 75 A; 143 C; 122 G; 134 T; 0 U; 0 Other;

  Query Match            7.6%;  Score 135.4;  DB 5;  Length 474;
  Best Local Similarity  72.6%;  Pred. No. 2.8e-26;
  Matches  175;  Conservative    0;  Mismatches   66;  Indels    0;  Gaps    0;

Qy         81 CTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGCAGGGTCACCTGTCTGTGTGC 140
              || |||   ||   | ||       || ||   |      ||||||||||||||||||||
Db        367 CTCTGGCCAGAAGCAAGAAGAAAGTTCTGACCACTGGCAGAGGGTCACCTGTCTGTGTGC 308

Qy        141 CCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGTCGAGGATTGGCACTCACAGC 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGTCGAGGATTGGCACTCACAGC 248

Qy        201 GGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACCGAAACAGAGCCCGTGTTCTG 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||      || |
Db        247 GGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACCGAAACAGGTGAACTGCGCCA 188

Qy        261 GTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAGCGCTTCTACTCCCTGCGCCA 320
              | |  |||     | |   ||  ||| | |||||| ||       ||  |||| | |
Db        187 GGCCACTGTGGCCATGATGAACAGGAAGGATGAGCTGGAGGAGGAGAACAGCCTGGGTGA 128

Qy        321 C 321
              |
Db        127 C 127


RESULT 15
AAC11927
ID   AAC11927 standard; cDNA; 190 BP.
XX
AC   AAC11927;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 16002.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
```

OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-00200610.
XX
PR 26-FEB-1999; 99US-0122487P.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS Claim 1; SEQ ID NO 16002; 71pp + Sequence Listing; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors
XX
SQ Sequence 190 BP; 47 A; 42 C; 73 G; 28 T; 0 U; 0 Other;

Query Match 7.5%; Score 133.4; DB 3; Length 190;
Best Local Similarity 99.3%; Pred. No. 6.2e-26;
Matches 134; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1464 CAGATCACTGCGAAACCTGCTCGACGGTGAGATGGAGCACTCAGCCGCGCTCCGGCAAGA 1523
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 CAGATCACTGCGAAACCTGCTCGACGGTGAGATGGAGCACTCAGCCGCGCTCCGGCAAGA 78

Qy       1524 GGTGGACACCTTGAAAAGGAAGGTGGCTGAACAGGAGGAGCGGCAGGGCATGAAGGTCCA 1583
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 GGTGGACACCTTGAAAAGGAAGGTGGCTGAACAGGAGGAGCGGCAGGGCATGAAGGTCCA 138

Qy       1584 GGCGCTGGCCAGCTA 1598
             ||||||||||||| ||
Db        139 GGCGCTGGCCAGGTA 153
```

Search completed: October 26, 2005, 02:00:13
Job time : 987 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 25, 2005, 21:20:38 ; Search time 7722 Seconds
(without alignments)
11181.973 Million cell updates/sec

Title: US-10-070-255-4
Perfect score: 1782
Sequence: 1 atgagcggatcacagaacaa.....taaattgctttcattccttg 1782

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1782	100.0	1782	6	AX089604	AX089604 Sequence
2	1742	97.8	1906	6	A66645	A66645 Sequence 1
3	1595	89.5	2873	6	AX089606	AX089606 Sequence
4	1595	89.5	3139	6	AX089605	AX089605 Sequence
5	1129.8	63.4	2248	6	BD157657	BD157657 Primer fo
6	1129.8	63.4	2248	6	AX879111	AX879111 Sequence
7	1129.8	63.4	2248	9	AK022425	AK022425 Homo sapi
8	937.4	52.6	1033	6	CQ714292	CQ714292 Sequence
9	818.4	45.9	5833	9	HSM808428	BX648280 Homo sapi
10	816.8	45.8	6045	6	BD160227	BD160227 Primer fo
11	816.8	45.8	6045	6	AX883269	AX883269 Sequence
12	816.8	45.8	6045	9	AK023827	AK023827 Homo sapi
13	654.8	36.7	1229	9	BC041583	BC041583 Homo sapi
14	639.8	35.9	730	6	BD146965	BD146965 Primer fo
15	639.8	35.9	730	6	AX866903	AX866903 Sequence
c 16	388	21.8	161864	9	AC007601	AC007601 Homo sapi
c 17	388	21.8	174136	2	AC126766	AC126766 Homo sapi
18	388	21.8	198521	2	AC145723	AC145723 Homo sapi
19	388	21.8	201237	2	AC141081	AC141081 Homo sapi
20	388	21.8	202689	2	AC034281	AC034281 Homo sapi
c 21	373.8	21.0	198703	2	AC144590	AC144590 Gorilla g
22	365.8	20.5	203216	2	AC097328	AC097328 Pan trogl
c 23	364.2	20.4	132393	2	AC138881	AC138881 Homo sapi
24	364.2	20.4	140011	2	AC008864	AC008864 Homo sapi
c 25	364.2	20.4	158388	9	AC008740	AC008740 Homo sapi
26	364.2	20.4	168268	9	AC009124	AC009124 Homo sapi
27	364.2	20.4	171940	9	AC025279	AC025279 Homo sapi
c 28	364.2	20.4	174010	2	AC023814	AC023814 Homo sapi
29	364.2	20.4	178081	2	AC023463	AC023463 Homo sapi
c 30	364.2	20.4	194943	2	AC142201	AC142201 Homo sapi
c 31	364.2	20.4	200812	9	AC133555	AC133555 Homo sapi
32	364.2	20.4	233479	2	AC126758	AC126758 Homo sapi
33	362.6	20.3	156989	2	AC069176	AC069176 Homo sapi
34	305.4	17.1	1431	10	AF399755	AF399755 Mus muscu
35	275	15.4	5496	9	AK074072	AK074072 Homo sapi
c 36	254.4	14.3	198521	2	AC145723	AC145723 Homo sapi
c 37	254.4	14.3	202689	2	AC034281	AC034281 Homo sapi
c 38	254.2	14.3	203216	2	AC097328	AC097328 Pan trogl
39	245.6	13.8	148230	2	AC141548	AC141548 Rattus no
40	245.6	13.8	165881	2	AC133732	AC133732 Rattus no
41	245.6	13.8	263443	2	AC121644	AC121644 Rattus no
42	233	13.1	177955	2	AC131034	AC131034 Mus muscu
c 43	227.8	12.8	156989	2	AC069176	AC069176 Homo sapi
44	205.4	11.5	227	6	CQ731480	CQ731480 Sequence
45	205	11.5	2945	10	BC034114	BC034114 Mus muscu

ALIGNMENTS

RESULT 1
AX089604
LOCUS AX089604 1782 bp DNA linear PAT 21-MAR-2001
DEFINITION Sequence 4 from Patent WO0116314.
ACCESSION AX089604
VERSION AX089604.1 GI:13443796
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Wallach,D., Malinin,N., Sinha,I.W. and Leu,S.
TITLE Iren protein, its preparation and use
JOURNAL Patent: WO 0116314-A 4 08-MAR-2001;
YEDA RESEARCH AND DEVELOPMENT Co. LTD. (IL)
FEATURES Location/Qualifiers
source 1. .1782
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 100.0%; Score 1782; DB 6; Length 1782;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1782; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60

Qy         61 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 120

Qy        121 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 180

Qy        181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 240
```

```
Qy       241 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 300
Db       241 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 300

Qy       301 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 360
Db       301 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 360

Qy       361 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 420
Db       361 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 420

Qy       421 AGGCTGAGCACTTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 480
Db       421 AGGCTGAGCACTTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 480

Qy       481 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 540
Db       481 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 540

Qy       541 GATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 600
Db       541 GATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 600

Qy       601 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 660
Db       601 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 660

Qy       661 ATCACAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTG 720
Db       661 ATCACAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTG 720

Qy       721 CCTGTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAA 780
Db       721 CCTGTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAA 780

Qy       781 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 840
Db       781 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 840

Qy       841 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCT 900
Db       841 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCT 900

Qy       901 GTCAATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGC 960
Db       901 GTCAATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGC 960

Qy       961 AGCAACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTT 1020
Db       961 AGCAACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTT 1020

Qy      1021 GATGTGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAAC 1080
Db      1021 GATGTGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAAC 1080

Qy      1081 TCATCAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAAGCCACTGGAAGGGAAC 1140
Db      1081 TCATCAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAAGCCACTGGAAGGGAAC 1140

Qy      1141 ACCTGCCTCTCCCAGATGCACAGCTGGGCTCCGCTGAAGGTGCTGCACAATGACTCCGAC 1200
Db      1141 ACCTGCCTCTCCCAGATGCACAGCTGGGCTCCGCTGAAGGTGCTGCACAATGACTCCGAC 1200

Qy      1201 ATCCTCTTCCCTGTCAGTGGCGTGGGCTCCTACAGCCCAGCAGATGCCCCCCTCGGAAGC 1260
Db      1201 ATCCTCTTCCCTGTCAGTGGCGTGGGCTCCTACAGCCCAGCAGATGCCCCCCTCGGAAGC 1260

Qy      1261 CTGGAGAACGGGACAGGACCAGAGGACCACGTTCTCCCGGATCCTGGACTTCGGTACAGT 1320
Db      1261 CTGGAGAACGGGACAGGACCAGAGGACCACGTTCTCCCGGATCCTGGACTTCGGTACAGT 1320

Qy      1321 GTGGAAGCCAGCTCTCCAGGCCACGGAAGTCCTCTGAGCAGCCTGTTACCTTCTGCCTCA 1380
Db      1321 GTGGAAGCCAGCTCTCCAGGCCACGGAAGTCCTCTGAGCAGCCTGTTACCTTCTGCCTCA 1380

Qy      1381 GTGCCAGAGTCCATGACAATTAGTGAACTGCGCCAGGCCACTGTGGCCATGATGAACAGG 1440
Db      1381 GTGCCAGAGTCCATGACAATTAGTGAACTGCGCCAGGCCACTGTGGCCATGATGAACAGG 1440

Qy      1441 AAGGATGAGCTGGAGGAGGAGAACAGATCACTGCGAAACCTGCTCGACGGTGAGATGGAG 1500
Db      1441 AAGGATGAGCTGGAGGAGGAGAACAGATCACTGCGAAACCTGCTCGACGGTGAGATGGAG 1500

Qy      1501 CACTCAGCCGCGCTCCGGCAAGAGGTGGACACCTTGAAAAGGAAGGTGGCTGAACAGGAG 1560
Db      1501 CACTCAGCCGCGCTCCGGCAAGAGGTGGACACCTTGAAAAGGAAGGTGGCTGAACAGGAG 1560

Qy      1561 GAGCGGCAGGGCATGAAGGTCCAGGCGCTGGCCAGCTATCTTTGCTATTTTGTGAGGAGA 1620
Db      1561 GAGCGGCAGGGCATGAAGGTCCAGGCGCTGGCCAGCTATCTTTGCTATTTTGTGAGGAGA 1620

Qy      1621 TTCTAACCCCACGTGAGAACCATGTGGTGGAGAAATGGAGGGAGAGAGAAATCCAACAGT 1680
Db      1621 TTCTAACCCCACGTGAGAACCATGTGGTGGAGAAATGGAGGGAGAGAGAAATCCAACAGT 1680

Qy      1681 TCCTGATAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCCTGAAGCTGTGTTTCCTCTGG 1740
Db      1681 TCCTGATAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCCTGAAGCTGTGTTTCCTCTGG 1740

Qy      1741 ACTTTTCATGTATGTGAGCCAATAAATTGCTTTCATTCCTTG 1782
Db      1741 ACTTTTCATGTATGTGAGCCAATAAATTGCTTTCATTCCTTG 1782
```

RESULT 2
A66645

LOCUS	A66645 1906 bp DNA linear PAT 29-MAR-1999
DEFINITION	Sequence 1 from Patent WO9737016.
ACCESSION	A66645
VERSION	A66645.1 GI:4538138
KEYWORDS	.
SOURCE	unidentified
ORGANISM	unidentified unclassified.
REFERENCE	1 (bases 1 to 1906)
AUTHORS	Wallach,D., Malinin,N., Boldin,M., Kovalenko,A. and Mett,I.
TITLE	MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR PREPARATION AND USE
JOURNAL	Patent: WO 9737016-A 1 09-OCT-1997; YEDA RES & DEV (IL)
COMMENT	Other publication AU 2175597 19971022.
FEATURES	Location/Qualifiers
source	1. .1906 /organism="unidentified" /mol_type="unassigned DNA" /db_xref="taxon:32644"
ORIGIN	

Query Match 97.8%; Score 1742; DB 6; Length 1906;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1775; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

```
Qy         1 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60
Db       118 ATGAGCGGATCNCNGAACNATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 177

Qy        61 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 120
Db       178 GTGAAACAGTGCCAGATCCGCTTTNGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 237

Qy       121 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 180
Db       238 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 297
```

```
Qy       181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 240
Db       298 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 357

Qy       241 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 300
Db       358 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 417

Qy       301 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 360
Db       418 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 477

Qy       361 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 420
Db       478 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 537

Qy       421 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 480
Db       538 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 597

Qy       481 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 540
Db       598 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 657

Qy       541 GATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 600
Db       658 GATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 717

Qy       601 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 660
Db       718 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 777

Qy       661 ATCACAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTG 720
Db       778 ATCACAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGA-CCCTTG 836

Qy       721 CCTGTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAA 780
Db       837 CCTGTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAA 896

Qy       781 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 840
Db       897 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 956

Qy       841 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCT 900
Db       957 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCT 1016

Qy       901 GTCAATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGC 960
Db      1017 GTCAATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAA-TCAGAGC 1075

Qy       961 AGCAACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTT 1020
Db      1076 AGCAACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTT 1135

Qy      1021 GATGTGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAAC 1080
Db      1136 GATGTGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAAC 1195

Qy      1081 TCATCAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAAGCCACTGGAAGGGAAC 1140
Db      1196 TCATCAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAAGCCACTGGAAGGGAAC 1255

Qy      1141 ACCTGCCTCTCCCAGATGCACAGCTGGGCTCCGCTGAAGGTGCTGCACAATGACTCCGAC 1200
Db      1256 ACCTGCCTCTCCCAGATGCACAGCTGGGCTCCGCTGAAGGTGCTGCACAATGACTCCGAC 1315

Qy      1201 ATCCTCTTCCCTGTCAGTGGCGTGGGCTCCTACAGCCCAGCAGATGCCCCCCTCGGAAGC 1260
Db      1316 ATCCTCTTCCCTGTCAGTGGCGTGGGCTCCTACAGCCCAGCAGATGCCCCCCTCGGAAGC 1375

Qy      1261 CTGGAGAACGGGACAGGACCAGAGGACCACGTTCTCCCGGATCCTGGACTTCGGTACAGT 1320
Db      1376 CTGGAGAACGGGACAGGACCAGAGGACCACGTTCTCCCGGATCCTGGACTTCGGTACAGT 1435

Qy      1321 GTGGAAGCCAGCTCTCCAGGCCACGGAAGTCCTCTGAGCAGCCTGTTACCTTCTGCCTCA 1380
Db      1436 GTGGAAGCCAGCTCTCCAGGCCACGGAAGTCCTCTGAGCAGCCTGTTA-CTTCTGCCTCA 1494

Qy      1381 GTGCCAGAGTCCATGACAATTAGTGAACTGCGCCAGGCCACTGTGGCCATGATGAACAGG 1440
Db      1495 GTGCCAGAGTCCATGACAATTAGTGAACTGCGCCAGGCCACTGTGGCCATGATGAACAGG 1554

Qy      1441 AAGGATGAGCTGGAGGAGGAGAACAGATCACTGCGAAACCTGCTCGACGGTGAGATGGAG 1500
Db      1555 AAGGATGAGCTGGAGGAGGAGAACAGATCACTGCGAAACCTGCTCGACGGTGAGATGGAG 1614

Qy      1501 CACTCAGCCGCGCTCCGGCAAGAGGTGGACACCTTGAAAAGGAAGGTGGCTGAACAGGAG 1560
Db      1615 CACTCAGCCGCGCTCCGGCAAGAGGTGGACACCTTGAAAAGGAAGGTGGCTGAACAGGAG 1674

Qy      1561 GAGCGGCAGGGCATGAAGGTCCAGGCGCTGGCCAGCTATCTTTGCTATTTTGTGAGGAGA 1620
Db      1675 GAGCGGCAGGGCATGAAGGTCCAGGCGCTGGCCAGCTATCTTTGCTATTTTGTGAGGAGA 1734

Qy      1621 TTCTAACCCCACGTGAGAACCATGTGGTGGAGAAATGGAGGGAGAGAGAAATCCAACAGT 1680
Db      1735 TTCTAACCCCACGTGAGAACCATGTGGTGGAGAAATGGAGGGAGAGAGAAATCCAACAGT 1794

Qy      1681 TCCTGATAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCCTGAAGCTGTGTTTCCTCTGG 1740
Db      1795 TCCTGATAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCCTGAAGCTGTGTTTCCTCTGG 1854

Qy      1741 ACTTTTCATGTATGTGAGCCAATAAATTGCTTTCATTCCTTG 1782
Db      1855 ACTTTTCATGTATGTGAGCCAATAAATTGCTTTCATTCCTTG 1896
```

RESULT 3
AX089606
LOCUS AX089606 2873 bp DNA linear PAT 21-MAR-2001
DEFINITION Sequence 6 from Patent WO0116314.
ACCESSION AX089606
VERSION AX089606.1 GI:13443798
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Wallach,D., Malinin,N., Sinha,I.W. and Leu,S.
TITLE Iren protein, its preparation and use
JOURNAL Patent: WO 0116314-A 6 08-MAR-2001;
YEDA RESEARCH AND DEVELOPMENT Co. LTD. (IL)
FEATURES Location/Qualifiers
source 1. .2873
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 89.5%; Score 1595; DB 6; Length 2873;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1595; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60
Db          1 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60

Qy         61 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 120
Db         61 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 120

Qy        121 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 180
```

```
Db        121 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 180
Qy        181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 240
Db        181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 240
Qy        241 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 300
Db        241 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 300
Qy        301 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 360
Db        301 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 360
Qy        361 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 420
Db        361 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 420
Qy        421 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 480
Db        421 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 480
Qy        481 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 540
Db        481 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 540
Qy        541 GATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 600
Db        541 GATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 600
Qy        601 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 660
Db        601 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 660
Qy        661 ATCACAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTG 720
Db        661 ATCACAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTG 720
Qy        721 CCTGTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAA 780
Db        721 CCTGTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAA 780
Qy        781 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 840
Db        781 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 840
Qy        841 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCT 900
Db        841 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCT 900
Qy        901 GTCAATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGC 960
Db        901 GTCAATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGC 960
Qy        961 AGCAACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTT 1020
Db        961 AGCAACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTT 1020
Qy       1021 GATGTGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAAC 1080
Db       1021 GATGTGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAAC 1080
Qy       1081 TCATCAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAAGCCACTGGAAGGGAAC 1140
Db       1081 TCATCAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAAGCCACTGGAAGGGAAC 1140
Qy       1141 ACCTGCCTCTCCCAGATGCACAGCTGGGCTCCGCTGAAGGTGCTGCACAATGACTCCGAC 1200
Db       1141 ACCTGCCTCTCCCAGATGCACAGCTGGGCTCCGCTGAAGGTGCTGCACAATGACTCCGAC 1200
Qy       1201 ATCCTCTTCCCTGTCAGTGGCGTGGGCTCCTACAGCCCAGCAGATGCCCCCCTCGGAAGC 1260
Db       1201 ATCCTCTTCCCTGTCAGTGGCGTGGGCTCCTACAGCCCAGCAGATGCCCCCCTCGGAAGC 1260
```

```
Qy       1261 CTGGAGAACGGGACAGGACCAGAGGACCACGTTCTCCCGGATCCTGGACTTCGGTACAGT 1320
Db       1261 CTGGAGAACGGGACAGGACCAGAGGACCACGTTCTCCCGGATCCTGGACTTCGGTACAGT 1320
Qy       1321 GTGGAAGCCAGCTCTCCAGGCCACGGAAGTCCTCTGAGCAGCCTGTTACCTTCTGCCTCA 1380
Db       1321 GTGGAAGCCAGCTCTCCAGGCCACGGAAGTCCTCTGAGCAGCCTGTTACCTTCTGCCTCA 1380
Qy       1381 GTGCCAGAGTCCATGACAATTAGTGAACTGCGCCAGGCCACTGTGGCCATGATGAACAGG 1440
Db       1381 GTGCCAGAGTCCATGACAATTAGTGAACTGCGCCAGGCCACTGTGGCCATGATGAACAGG 1440
Qy       1441 AAGGATGAGCTGGAGGAGGAGAACAGATCACTGCGAAACCTGCTCGACGGTGAGATGGAG 1500
Db       1441 AAGGATGAGCTGGAGGAGGAGAACAGATCACTGCGAAACCTGCTCGACGGTGAGATGGAG 1500
Qy       1501 CACTCAGCCGCGCTCCGGCAAGAGGTGGACACCTTGAAAAGGAAGGTGGCTGAACAGGAG 1560
Db       1501 CACTCAGCCGCGCTCCGGCAAGAGGTGGACACCTTGAAAAGGAAGGTGGCTGAACAGGAG 1560
Qy       1561 GAGCGGCAGGGCATGAAGGTCCAGGCGCTGGCCAG 1595
Db       1561 GAGCGGCAGGGCATGAAGGTCCAGGCGCTGGCCAG 1595
```

RESULT 4
AX089605
LOCUS AX089605 3139 bp DNA linear PAT 21-MAR-2001
DEFINITION Sequence 5 from Patent WO0116314.
ACCESSION AX089605
VERSION AX089605.1 GI:13443797
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Wallach,D., Malinin,N., Sinha,I.W. and Leu,S.
TITLE Iren protein, its preparation and use
JOURNAL Patent: WO 0116314-A 5 08-MAR-2001;
YEDA RESEARCH AND DEVELOPMENT Co. LTD. (IL)
FEATURES Location/Qualifiers
source 1. .3139
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 89.5%; Score 1595; DB 6; Length 3139;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1595; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60
Db          1 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60
Qy         61 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 120
Db         61 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 120
Qy        121 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 180
Db        121 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 180
Qy        181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 240
Db        181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 240
Qy        241 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 300
Db        241 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 300
```

```
Qy        301 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 360

Qy        361 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 420

Qy        421 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 480

Qy        481 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 540

Qy        541 GATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 600

Qy        601 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 660

Qy        661 ATCACAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATCACAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTG 720

Qy        721 CCTGTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTGTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAA 780

Qy        781 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 840

Qy        841 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCT 900

Qy        901 GTCAATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GTCAATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGC 960

Qy        961 AGCAACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGCAACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTT 1020

Qy       1021 GATGTGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GATGTGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAAC 1080

Qy       1081 TCATCAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAAGCCACTGGAAGGGAAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCATCAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAAGCCACTGGAAGGGAAC 1140

Qy       1141 ACCTGCCTCTCCCAGATGCACAGCTGGGCTCCGCTGAAGGTGCTGCACAATGACTCCGAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACCTGCCTCTCCCAGATGCACAGCTGGGCTCCGCTGAAGGTGCTGCACAATGACTCCGAC 1200

Qy       1201 ATCCTCTTCCCTGTCAGTGGCGTGGGCTCCTACAGCCCAGCAGATGCCCCCCTCGGAAGC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATCCTCTTCCCTGTCAGTGGCGTGGGCTCCTACAGCCCAGCAGATGCCCCCCTCGGAAGC 1260

Qy       1261 CTGGAGAACGGGACAGGACCAGAGGACCACGTTCTCCCGGATCCTGGACTTCGGTACAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTGGAGAACGGGACAGGACCAGAGGACCACGTTCTCCCGGATCCTGGACTTCGGTACAGT 1320

Qy       1321 GTGGAAGCCAGCTCTCCAGGCCACGGAAGTCCTCTGAGCAGCCTGTTACCTTCTGCCTCA 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTGGAAGCCAGCTCTCCAGGCCACGGAAGTCCTCTGAGCAGCCTGTTACCTTCTGCCTCA 1380

Qy       1381 GTGCCAGAGTCCATGACAATTAGTGAACTGCGCCAGGCCACTGTGGCCATGATGAACAGG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GTGCCAGAGTCCATGACAATTAGTGAACTGCGCCAGGCCACTGTGGCCATGATGAACAGG 1440

Qy       1441 AAGGATGAGCTGGAGGAGGAGAACAGATCACTGCGAAACCTGCTCGACGGTGAGATGGAG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AAGGATGAGCTGGAGGAGGAGAACAGATCACTGCGAAACCTGCTCGACGGTGAGATGGAG 1500

Qy       1501 CACTCAGCCGCGCTCCGGCAAGAGGTGGACACCTTGAAAAGGAAGGTGGCTGAACAGGAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CACTCAGCCGCGCTCCGGCAAGAGGTGGACACCTTGAAAAGGAAGGTGGCTGAACAGGAG 1560

Qy       1561 GAGCGGCAGGGCATGAAGGTCCAGGCGCTGGCCAG 1595
              ||||||||||||||||||||||||||||||||||||
Db       1561 GAGCGGCAGGGCATGAAGGTCCAGGCGCTGGCCAG 1595
```

RESULT 5
BD157657

```
LOCUS       BD157657               2248 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Primer for synthesizing full-length cDNA and use thereof.
ACCESSION   BD157657
VERSION     BD157657.1  GI:27863415
KEYWORDS    JP 2002191363-A/12500.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2248)
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primer for synthesizing full-length cDNA and use thereof
  JOURNAL   Patent: JP 2002191363-A 12500 09-JUL-2002;
            HELIX RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002191363-A/12500
            PD   09-JUL-2002
            PF   28-JUL-2000 JP 2000280990
            PI   TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
            PI   SAITO,
            PI   JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
            PI   KEIICHI NAGAI,TETSUJI OTSUKI
            PC
            C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
             10,
            PC   C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
             Primer for synthesizing full-length cDNA and use thereof FH   Key
                      Location/Qualifiers
            FT   CDS             (63). .(1187).
FEATURES             Location/Qualifiers
     source          1. .2248
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

Query Match 63.4%; Score 1129.8; DB 6; Length 2248;
Best Local Similarity 95.3%; Pred. No. 3.8e-267;
Matches 1164; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

```
Qy          1 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 122

Qy         61 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 182

Qy        121 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 242

Qy        181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        243 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 302

Qy        241 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 362

Qy        301 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 422

Qy        361 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 482

Qy        421 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 542

Qy        481 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 602

Qy        541 GATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 600
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 GATTTGGACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 662

Qy        601 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 722

Qy        661 ATCACAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 ATCACAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTG 782

Qy        721 CCTGTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAA 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 CCTGTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAG 842

Qy        781 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        843 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 902

Qy        841 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        903 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCT 962

Qy        901 GTCAATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 GTCAATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGC 1022

Qy        961 AGCAACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 AGCAACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTT 1082

Qy       1021 GATGTGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAAC 1080
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1083 GATGTGAAAAGGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAAC 1142

Qy       1081 TCATCAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAAGCCACTGGAAGGGAAC 1140
              |||||||||||||||||||||||||||||||||||||||||||  | | |    | | |
Db       1143 TCATCAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAAGTAAGTTTGTGTGTAA 1202

Qy       1141 ACCTGCCTCTCCCAGATGCACAGCTGGGCTCCGCTGAAGGTGCTGCACAATGACTCCGAC 1200
                  | |||| | |    | |||   ||      |     ||| |       | ||| | ||
Db       1203 GGTGGAGTCTCACCGTCCCCCAGGCTGGAGTGCCGTGGCGTGATCTCAGCTCACTGCAAC 1262

Qy       1201 ATCCTCTTCCCTGTCAGTGGC 1221
                |||| | |||| |    |||
Db       1263 CTCCACCTCCCGGGTTCAGGC 1283


RESULT 6
AX879111
```

```
LOCUS       AX879111                2248 bp    DNA     linear   PAT 17-DEC-2003
DEFINITION  Sequence 14016 from Patent EP1074617.
ACCESSION   AX879111
VERSION     AX879111.1  GI:40033847
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primers for synthesising full-length cDNA and their use
  JOURNAL   Patent: EP 1074617-A 14016 07-FEB-2001;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .2248
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             63. .1190
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAE90370.1"
                     /db_xref="GI:40033848"
                     /translation="MSGSQNNDKRQFLLERLLDAVKQCQIRFGGRKEIASDSDSRVTC
                     LCAQFEAVLQHGLKRSRGLALTAAAIKQAAGFASKTETEPVFWYYVKEVLNKHELQRF
                     YSLRHIASDVGRGRAWLRCALNEHSLERYLHMLLADRCRLSTFYEDWSFVMDEERSSM
                     LPTMAAGLNSILFAINIDNKDLDGQSKFAPTVSDLLKESTQNVTSLLKESTQGVSSLF
                     REITASSAVSILIKPEQETDPLPVVSRNVSADAKCKKERKKKKKVTNIISFDDEEDEQ
                     NSGDVFKKTPGAGESSEDNSDRSSVNIMSAFESPFGPNSNGSQSSNSWKIDSLSLNGE
                     FGYQKLDVKGIDDEDVDENEDDVYGNSSGRKHRGHSESPEK"
ORIGIN

  Query Match           63.4%;  Score 1129.8;  DB 6;  Length 2248;
  Best Local Similarity  95.3%;  Pred. No. 3.8e-267;
  Matches 1164;  Conservative    0;  Mismatches   57;  Indels    0;  Gaps    0;

Qy          1 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 122

Qy         61 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 182

Qy        121 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 242

Qy        181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 302

Qy        241 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 362

Qy        301 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 422

Qy        361 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 482

Qy        421 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 542

Qy        481 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 602
```

```
Qy        541 GATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 600
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 GATTTGGACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 662

Qy        601 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 722

Qy        661 ATCACAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 ATCACAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTG 782

Qy        721 CCTGTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 CCTGTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAG 842

Qy        781 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        843 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 902

Qy        841 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        903 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCT 962

Qy        901 GTCAATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 GTCAATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGC 1022

Qy        961 AGCAACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 AGCAACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTT 1082

Qy       1021 GATGTGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAAC 1080
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1083 GATGTGAAAGGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAAC 1142

Qy       1081 TCATCAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAAGCCACTGGAAGGGAAC 1140
              |||||||||||||||||||||||||||||||||||||||||||  | |    | | |
Db       1143 TCATCAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAAGTAAGTTTGTGTGTAA 1202

Qy       1141 ACCTGCCTCTCCCAGATGCACAGCTGGGCTCCGCTGAAGGTGCTGCACAATGACTCCGAC 1200
                  |||| | |   | |||   ||      |     ||| |     | ||| | ||
Db       1203 GGTGGAGTCTCACCGTCCCCCAGGCTGGAGTGCCGTGGCGTGATCTCAGCTCACTGCAAC 1262

Qy       1201 ATCCTCTTCCCTGTCAGTGGC 1221
                ||| | |||| |    |||
Db       1263 CTCCACCTCCCGGGTTCAGGC 1283
```

RESULT 7
AK022425
LOCUS AK022425 2248 bp mRNA linear PRI 30-JAN-2004
DEFINITION Homo sapiens cDNA FLJ12363 fis, clone MAMMA1002380.
ACCESSION AK022425
VERSION AK022425.1 GI:10433818
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R., Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H., Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y., Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M., Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T., Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H., Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M., Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S., Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A., Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A., Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T., Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S., Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H., Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T., Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M., Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T., Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J., Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K., Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R., Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.
TITLE Complete sequencing and characterization of 21,243 full-length human cDNAs
JOURNAL Nat. Genet. 36 (1), 40-45 (2004)
PUBMED 14702039
REFERENCE 2
AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H., Wagatsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K., Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 2248)
AUTHORS Isogai,T. and Otsuki,T.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
FEATURES Location/Qualifiers
source 1. .2248
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MAMMA1002380"
/tissue_type="Mammary gland"
/clone_lib="MAMMA1"
/note="cloning vector: pME18SFL3"
CDS 63. .1190
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB14033.1"
/db_xref="GI:10433819"
/translation="MSGSQNNDKRQFLLERLLDAVKQCQIRFGGRKEIASDSDSRVTC
LCAQFEAVLQHGLKRSRGLALTAAAIKQAAGFASKTETEPVFWYYVKEVLNKHELQRF
YSLRHIASDVGRGRAWLRCALNEHSLERYLHMLLADRCRLSTFYEDWSFVMDEERSSM
LPTMAAGLNSILFAINIDNKDLDGQSKFAPTVSDLLKESTQNVTSLLKESTQGVSSLF
REITASSAVSILIKPEQETDPLPVVSRNVSADAKCKKERKKKKKVTNIISFDDEEDEQ
NSGDVFKKTPGAGESSEDNSDRSSVNIMSAFESPFGPNSNGSQSSNSWKIDSLSLNGE
FGYQKLDVKGIDDEDVDENEDDVYGNSSGRKHRGHSESPEK"
ORIGIN

Query Match 63.4%; Score 1129.8; DB 9; Length 2248;
Best Local Similarity 95.3%; Pred. No. 3.8e-267;
Matches 1164; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

```
Qy          1 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60
```

```
Db         63 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 122

Qy         61 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 120

Db        123 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 182

Qy        121 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 180

Db        183 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 242

Qy        181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 240

Db        243 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 302

Qy        241 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 300

Db        303 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 362

Qy        301 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 360

Db        363 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 422

Qy        361 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 420

Db        423 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 482

Qy        421 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 480

Db        483 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 542

Qy        481 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 540

Db        543 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 602

Qy        541 GATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 600

Db        603 GATTTGGACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 662

Qy        601 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 660

Db        663 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 722

Qy        661 ATCACAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTG 720

Db        723 ATCACAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTG 782

Qy        721 CCTGTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAA 780

Db        783 CCTGTCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAG 842

Qy        781 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 840

Db        843 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 902

Qy        841 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCT 900

Db        903 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCT 962

Qy        901 GTCAATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGC 960

Db        963 GTCAATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGC 1022

Qy        961 AGCAACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTT 1020

Db       1023 AGCAACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTT 1082

Qy       1021 GATGTGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAAC 1080

Db       1083 GATGTGAAAGGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAAC 1142

Qy       1081 TCATCAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAAGCCACTGGAAGGGAAC 1140

Db       1143 TCATCAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAAGTAAGTTTGTGTGTAA 1202

Qy       1141 ACCTGCCTCTCCCAGATGCACAGCTGGGCTCCGCTGAAGGTGCTGCACAATGACTCCGAC 1200

Db       1203 GGTGGAGTCTCACCGTCCCCCAGGCTGGAGTGCCGTGGCGTGATCTCAGCTCACTGCAAC 1262

Qy       1201 ATCCTCTTCCCTGTCAGTGGC 1221

Db       1263 CTCCACCTCCCGGGTTCAGGC 1283
```

RESULT 8
CQ714292

```
LOCUS       CQ714292               1033 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 226 from Patent WO02068579.
ACCESSION   CQ714292
VERSION     CQ714292.1  GI:42275149
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 226 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .1033
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

Query Match 52.6%; Score 937.4; DB 6; Length 1033;
Best Local Similarity 92.4%; Pred. No. 9.5e-220;
Matches 1032; Conservative 0; Mismatches 1; Indels 84; Gaps 1;

```
Qy          8 GATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCAGTGAAAC 67

Db          1 GATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCAGTGAAAC 60

Qy         68 AGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGCAGGGTCA 127

Db         61 AGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGCAGGGTCA 120

Qy        128 CCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGTCGAGGAT 187

Db        121 CCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGTCGAGGAT 180

Qy        188 TGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACCGAAACAG 247

Db        181 TGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACCGAAA--- 237

Qy        248 AGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAGCGCTTCT 307

Db        238 ------------------------------------------------------------ 237

Qy        308 ACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGCTGTGCCC 367

Db        238 --------------------CAGACGTGGGCCGGGGCCGCGCCTGGCTGCGCTGTGCCC 276

Qy        368 TCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGCAGGCTGA 427

Db        277 TCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGCAGGCTGA 336

Qy        428 GCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTTCCTA 487

Db        337 GCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTTCCTA 396

Qy        488 CCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAGGATTTGA 547
```

```
Db        397 CCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAGGATTTGA 456

Qy        548 ACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAACG 607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 ACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAACG 516

Qy        608 TGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCACAG 667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        517 TGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCACAG 576

Qy        668 CCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCTGTCG 727
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        577 CCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCTGTCG 636

Qy        728 TGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGAAAG 787
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        637 TGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGAAAG 696

Qy        788 TGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGTTTA 847
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        697 TGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGTTTA 756

Qy        848 AAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCAATA 907
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        757 AAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCAATA 816

Qy        908 TCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGCAGCAACT 967
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        817 TCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGCAGCAACT 876

Qy        968 CATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTTGATGTGA 1027
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        877 CATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTTGATGTGA 936

Qy       1028 AAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCATCAG 1087
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        937 AAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCATCAG 996

Qy       1088 GAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAA 1124
              |||||||||||||||||||||||||||||||||||
Db        997 GAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAA 1033


RESULT 9
HSM808428
LOCUS       HSM808428              5833 bp    mRNA    linear   PRI 30-AUG-2003
DEFINITION  Homo sapiens mRNA; cDNA DKFZp686E09125 (from clone DKFZp686E09125).
ACCESSION   BX648280
VERSION     BX648280.1  GI:34367439
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5833)
  AUTHORS   Bloecker,H., Boecher,M., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
  CONSRTM   The German Human cDNA Consortium
  TITLE     Direct Submission
  JOURNAL   Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
            Neuherberg, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by GBF (National Research Centre for Biotechnology Ltd.,
            Braunschweig/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            This clone (DKFZp686E09125) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://mips.gsf.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1. .5833
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686E09125"
                     /tissue_type="human uterus"
                     /clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
                     DH10B; sites SfiIA + SfiIB"
                     /dev_stage="adult"
     polyA_signal    5789. .5794
     polyA_site      5810
ORIGIN

  Query Match             45.9%;  Score 818.4;  DB 9;  Length 5833;
  Best Local Similarity   97.2%;  Pred. No. 2.1e-190;
  Matches  855;  Conservative    0;  Mismatches   21;  Indels    4;  Gaps    2;

Qy        246 AGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGC-AGCGCT 304
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db       1381 AGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAAGCGCT 1440

Qy        305 TCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGCTGTG 364
              |||||||||||||||||| |||||||||||||||||||| ||| ||||||||||||||||
Db       1441 TCTACTCCCTGCGCCACATTGCCTCAGACGTGGGCCGGGGCCGCACCTGGCTGCGCTGTG 1500

Qy        365 CCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGCAGGC 424
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGCAGGC 1560

Qy        425 TGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTTC 484
              | || |||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db       1561 TCAGTACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAGAGGTCCAGTATGCTTC 1620

Qy        485 CTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAGGATT 544
              |||||||||||||||||| |||||||||||||||||||||||||| ||||||||||||||
Db       1621 CTACCATGGCAGCAGGTCCGAACTCCATACTCTTTGCGATTAACATTGACAACAAGGATT 1680

Qy        545 TGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGA 604
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGA 1740

Qy        605 ACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCA 664
              | ||||   |||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       1741 ATGTGA---CCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCGTGTTCAGGGAGATCA 1797

Qy        665 CAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCTG 724
              |||||||||||||| |||||||||||||||||||||||||||||||||||||||||| |
Db       1798 CAGCCTCCTCTGCCATCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCCG 1857

Qy        725 TCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGA 784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1858 TCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGC 1917

Qy        785 AAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGT 844
              |||||||||||||| |||||||||||||||||||||||||||||||||||| ||| |||
Db       1918 AAGTGACCAACATTATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACATGT 1977

Qy        845 TTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCA 904
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       1978 TTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCACTCCTCTGTCA 2037

Qy        905 ATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGCAGCA 964
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db       2038 ATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCAAACTCCAATGGAAGTCAGAGCAGCA 2097

Qy        965 ACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTTGATG 1024
              |||| |||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       2098 ACTCGTGGAAAATTGATTCCCTGTCTTTGAACAGGGAGTTTGGGTACCAGAAGCTTGATG 2157

Qy       1025 TGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCAT 1084
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2158 TGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCAT 2217
```

```
Qy       1085 CAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAA 1124
              ||||| |||||||||||||||||| |||||||||||||
Db       2218 CAGGACGGAAGCACAGGGGCCACTCAGAGTCGCCCGAGAA 2257


RESULT 10
BD160227
LOCUS       BD160227                6045 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Primer for synthesizing full-length cDNA and use thereof.
ACCESSION   BD160227
VERSION     BD160227.1  GI:27865985
KEYWORDS    JP 2002191363-A/15070.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 6045)
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primer for synthesizing full-length cDNA and use thereof
  JOURNAL   Patent: JP 2002191363-A 15070 09-JUL-2002;
            HELIX RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002191363-A/15070
            PD   09-JUL-2002
            PF   28-JUL-2000 JP 2000280990
            PI   TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
            PI   SAITO,
            PI   JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
            PI   KEIICHI NAGAI,TETSUJI OTSUKI
            PC
            C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
             10,
            PC   C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
             Primer for synthesizing full-length cDNA and use thereof FH   Key
                     Location/Qualifiers
            FT   CDS             (364). .(954).
FEATURES             Location/Qualifiers
     source          1. .6045
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           45.8%;  Score 816.8;  DB 6;  Length 6045;
  Best Local Similarity   97.0%;  Pred. No. 5.2e-190;
  Matches  854;  Conservative    0;  Mismatches   22;  Indels    4;  Gaps    2;

Qy        246 AGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGC-AGCGCT 304
              | |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        932 AAAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAAGCGCT 991

Qy        305 TCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGCTGTG 364
              |||||||||||||||||||||||||||||||||||||| ||| ||||||||||||||||
Db        992 TCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGCCGCACCTGGCTGCGCTGTG 1051

Qy        365 CCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGCAGGC 424
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1052 CCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGCAGGC 1111

Qy        425 TGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTTC 484
              | || |||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db       1112 TCAGTACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAGAGGTCCAGTATGCTTC 1171

Qy        485 CTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAGGATT 544
              |||||||||||||||| |||||||||||||||||||||| ||||||||||||||||||||
Db       1172 CTACCATGGCAGCAGGTCCGAACTCCATACTCTTTGCGATTAACATTGACAACAAGGATT 1231

Qy        545 TGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGA 604
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1232 TGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGA 1291

Qy        605 ACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCA 664
              | ||||   |||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       1292 ATGTGA---CCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCGTGTTCAGGGAGATCA 1348

Qy        665 CAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCTG 724
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||| |
Db       1349 CAGCCTCCTCTGCCATCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCCG 1408

Qy        725 TCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGA 784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1409 TCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGC 1468

Qy        785 AAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGT 844
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||| |||
Db       1469 AAGTGACCAACATTATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACATGT 1528

Qy        845 TTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCA 904
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       1529 TTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCACTCCTCTGTCA 1588

Qy        905 ATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGCAGCA 964
              |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db       1589 ATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCGAAACTCCAATGGAAGTCAGAGCAGCA 1648

Qy        965 ACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTTGATG 1024
              |||| |||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       1649 ACTCGTGGAAAATTGATTCCCTGTCTTTGAACAGGGAGTTTGGGTACCAGAAGCTTGATG 1708

Qy       1025 TGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCAT 1084
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1709 TGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCAT 1768

Qy       1085 CAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAA 1124
              ||||| |||||||||||||||||| |||||||||||||
Db       1769 CAGGACGGAAGCACAGGGGCCACTCAGAGTCGCCCGAGAA 1808


RESULT 11
AX883269
LOCUS       AX883269                6045 bp    DNA     linear   PAT 17-DEC-2003
DEFINITION  Sequence 18174 from Patent EP1074617.
ACCESSION   AX883269
VERSION     AX883269.1  GI:40038170
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primers for synthesising full-length cDNA and their use
  JOURNAL   Patent: EP 1074617-A 18174 07-FEB-2001;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .6045
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             364. .957
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAE91902.1"
                     /db_xref="GI:40038171"
                     /translation="MRVHCAIIPSDMLHISTNCRTAEEMVLTLLDYLFHHEVQAVSSL
                     SGQGKHGKKQLDPLTIYGIRCHLFYKFGITESDWYRIKQSIDSKCRTAWRRKQWGQSL
                     VVKSFSRRMPNSSSYCPSEPMMSTPPPARELLQPQLQPQALHYALANVQQVQIHQIGE
                     DGQVQVIPQGHLHIAKVPQGEQVQITQDSESPCSGTT"
ORIGIN

  Query Match           45.8%;  Score 816.8;  DB 6;  Length 6045;
  Best Local Similarity   97.0%;  Pred. No. 5.2e-190;
  Matches  854;  Conservative    0;  Mismatches   22;  Indels    4;  Gaps    2;
```

```
Qy        246 AGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGC-AGCGCT 304
              | ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        932 AAAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAAGCGCT 991

Qy        305 TCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGCTGTG 364
              |||||||||||||||||||||||||||||||||||||| ||| ||||||||||||||||
Db        992 TCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGCCGCACCTGGCTGCGCTGTG 1051

Qy        365 CCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGCAGGC 424
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1052 CCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGCAGGC 1111

Qy        425 TGAGCACTTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTTC 484
              | || ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       1112 TCAGTACTTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAGAGGTCCAGTATGCTTC 1171

Qy        485 CTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAGGATT 544
              |||||||||||||||||| |||||||||||||||||||||||||| ||||||||||||||
Db       1172 CTACCATGGCAGCAGGTCCGAACTCCATACTCTTTGCGATTAACATTGACAACAAGGATT 1231

Qy        545 TGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGA 604
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1232 TGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGA 1291

Qy        605 ACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCA 664
              | ||||   ||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       1292 ATGTGA---CCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCGTGTTCAGGGAGATCA 1348

Qy        665 CAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCTG 724
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||| |
Db       1349 CAGCCTCCTCTGCCATCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCCG 1408

Qy        725 TCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGA 784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1409 TCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGC 1468

Qy        785 AAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGT 844
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Db       1469 AAGTGACCAACATTATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACATGT 1528

Qy        845 TTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCA 904
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       1529 TTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCACTCCTCTGTCA 1588

Qy        905 ATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGCAGCA 964
              |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db       1589 ATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCGAAACTCCAATGGAAGTCAGAGCAGCA 1648

Qy        965 ACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTTGATG 1024
              |||| |||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       1649 ACTCGTGGAAAATTGATTCCCTGTCTTTGAACAGGGAGTTTGGGTACCAGAAGCTTGATG 1708

Qy       1025 TGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCAT 1084
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1709 TGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCAT 1768

Qy       1085 CAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAA 1124
              ||||| |||||||||||||||| ||||||||||||||
Db       1769 CAGGACGGAAGCACAGGGGCCACTCAGAGTCGCCCGAGAA 1808
```

```
RESULT 12
AK023827
LOCUS       AK023827               6045 bp    mRNA    linear   PRI 30-JAN-2004
DEFINITION  Homo sapiens cDNA FLJ13765 fis, clone PLACE4000128, weakly similar
            to Mus musculus putative transcription factor mRNA.
ACCESSION   AK023827
VERSION     AK023827.1  GI:10435879
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
            Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,
            Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
            Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y.,
            Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M.,
            Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
            Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T.,
            Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K.,
            Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H.,
            Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,
            Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S.,
            Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A.,
            Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R.,
            Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A.,
            Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T.,
            Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S.,
            Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H.,
            Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T.,
            Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K.,
            Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M.,
            Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y.,
            Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N.,
            Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,
            Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K.,
            Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T.,
            Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J.,
            Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K.,
            Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R.,
            Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.
  TITLE     Complete sequencing and characterization of 21,243 full-length
            human cDNAs
  JOURNAL   Nat. Genet. 36 (1), 40-45 (2004)
   PUBMED   14702039
REFERENCE   2
  AUTHORS   Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
            Nishikawa,T., Nagai,K., Sugano,S., Takahashi-Fujii,A., Hara,H.,
            Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K.,
            Arita,M., Nabekura,T., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J.,
            Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuho,Y. and Oshima,A.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 6045)
  AUTHORS   Isogai,T. and Otsuki,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,
            Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing and clone selection:
            Helix Research Institute (supported by Japan Key Technology Center
            etc.) and Department of Virology, Institute of Medical Science,
            University of Tokyo.
FEATURES             Location/Qualifiers
     source          1. .6045
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE4000128"
                     /tissue_type="placenta"
                     /clone_lib="PLACE4"
                     /note="cloning vector: pME18SFL3"
     CDS             364. .957
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAB14692.1"
                     /db_xref="GI:10435880"
                     /translation="MRVHCAIIPSDMLHISTNCRTAEEMVLTLLDYLFHHEVQAVSSL
                     SGQGKHGKKQLDPLTIYGIRCHLFYKFGITESDWYRIKQSIDSKCRTAWRRKQWGQSL
                     VVKSFSRRMPNSSSYCPSEPMMSTPPPARELLQPQLQPQALHYALANVQQVQIHQIGE
                     DGQVQVIPQGHLHIAKVPQGEQVQITQDSESPCSGTT"
```

```
ORIGIN

  Query Match           45.8%;  Score 816.8;  DB 9;  Length 6045;
  Best Local Similarity  97.0%;  Pred. No. 5.2e-190;
  Matches  854;  Conservative    0;  Mismatches   22;  Indels    4;  Gaps    2;

Qy        246 AGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGC-AGCGCT 304
              | |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        932 AAAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAAGCGCT 991

Qy        305 TCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGCTGTG 364
              |||||||||||||||||||||||||||||||||||||||||| ||| ||||||||||||||
Db        992 TCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGCCGCACCTGGCTGCGCTGTG 1051

Qy        365 CCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGCAGGC 424
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1052 CCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGCAGGC 1111

Qy        425 TGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTTC 484
              | ||| ||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db       1112 TCAGTACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAGAGGTCCAGTATGCTTC 1171

Qy        485 CTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAGGATT 544
              |||||||||||||||||| ||||||||||||||||||||||||||| ||||||||||||||
Db       1172 CTACCATGGCAGCAGGTCCGAACTCCATACTCTTTGCGATTAACATTGACAACAAGGATT 1231

Qy        545 TGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGA 604
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1232 TGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGA 1291

Qy        605 ACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCA 664
              | ||||   |||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       1292 ATGTGA---CCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCGTGTTCAGGGAGATCA 1348

Qy        665 CAGCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCTG 724
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||| |
Db       1349 CAGCCTCCTCTGCCATCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCCG 1408

Qy        725 TCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGA 784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1409 TCGTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGC 1468

Qy        785 AAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGT 844
              |||||||||||| |||||||||||||||||||||||||||||||||||||||||| |||
Db       1469 AAGTGACCAACATTATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACATGT 1528

Qy        845 TTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCA 904
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       1529 TTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCACTCCTCTGTCA 1588

Qy        905 ATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGCAGCA 964
              |||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||
Db       1589 ATATCATGTCCGCCTTTGAAAGCCCCTTCGGGCGAAACTCCAATGGAAGTCAGAGCAGCA 1648

Qy        965 ACTCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTTGATG 1024
              |||| |||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db       1649 ACTCGTGGAAAATTGATTCCCTGTCTTTGAACAGGGAGTTTGGGTACCAGAAGCTTGATG 1708

Qy       1025 TGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCAT 1084
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1709 TGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCAT 1768

Qy       1085 CAGGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAA 1124
              ||||| |||||||||||||||||| ||||||||||||||
Db       1769 CAGGACGGAAGCACAGGGGCCACTCAGAGTCGCCCGAGAA 1808


RESULT 13
BC041583
LOCUS       BC041583                1229 bp    mRNA    linear   PRI 07-OCT-2003
DEFINITION  Homo sapiens hypothetical protein FLJ12363, mRNA (cDNA clone
            MGC:52410 IMAGE:4765274), complete cds.
ACCESSION   BC041583
VERSION     BC041583.1  GI:27371337
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1229)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1229)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-DEC-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis Staudt
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 44 Row: e Column: 20.
FEATURES             Location/Qualifiers
     source          1. .1229
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:52410 IMAGE:4765274"
                     /tissue_type="Primary B-Cells from Tonsils"
                     /clone_lib="NIH_MGC_48"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     gene            1. .1229
                     /gene="FLJ12363"
                     /db_xref="LocusID:84127"
```

```
     CDS             182. .931
                     /gene="FLJ12363"
                     /codon_start=1
                     /product="FLJ12363 protein"
                     /protein_id="AAH41583.1"
                     /db_xref="GI:27371338"
                     /db_xref="LocusID:84127"
                     /translation="MDEERSSMLPTMAAGPNSILFAINIDNKDLNGQSKFAPTVSDLL
                     KESTQNVTLLKESTQGVSSVFREITASSAISILIKPEQETDPLPVVSRNVSADAKCKK
                     ERKKKKQVTNIISFDDEEDEQNSGDMFKKTPGAGESSEDNSDHSSVNIMSAFESPFGP
                     NSNGSQSSNSWKIDSLSLNREFGYQKLDVKSIDDEDVDENEDDVYGNSSGRKHRGHSE
                     SPEKNGAHSVTQAGVQWHDLGSLQPLPPGFR"
ORIGIN

  Query Match           36.7%;  Score 654.8;  DB 9;  Length 1229;
  Best Local Similarity   97.1%;  Pred. No. 4.1e-150;
  Matches  678;  Conservative    0;  Mismatches   17;  Indels    3;  Gaps    1;

Qy        427 AGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTTCCT 486
              || |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        152 AGTACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAGAGGTCCAGTATGCTTCCT 211

Qy        487 ACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAGGATTTG 546
              |||||||||||||||| |||||||||||||||||||||||||||| ||||||||||||||||||
Db        212 ACCATGGCAGCAGGTCCGAACTCCATACTCTTTGCGATTAACATTGACAACAAGGATTTG 271

Qy        547 AACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAAC 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 AACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGAAT 331

Qy        607 GTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCACA 666
              ||||   |||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        332 GTGA---CCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCGTGTTCAGGGAGATCACA 388

Qy        667 GCCTCCTCTGCCGTCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCTGTC 726
              |||||||||||| |||||||||||||||||||||||||||||||||||||||||||| |||
Db        389 GCCTCCTCTGCCATCTCCATCCTCATCAAACCTGAACAGGAGACCGACCCCTTGCCCGTC 448

Qy        727 GTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGAAA 786
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        449 GTGTCCAGGAATGTCAGTGCTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAAAGCAA 508

Qy        787 GTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGTTT 846
              |||||||||||| |||||||||||||||||||||||||||||||||||||||||||| |||||
Db        509 GTGACCAACATTATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACATGTTT 568

Qy        847 AAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCAAT 906
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        569 AAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCACTCCTCTGTCAAT 628

Qy        907 ATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGCAGCAAC 966
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        629 ATCATGTCCGCCTTTGAAAGCCCCTTCGGGCCAAACTCCAATGGAAGTCAGAGCAGCAAC 688

Qy        967 TCATGGAAAATTGATTCCCTGTCTTTGAACGGGGAGTTTGGGTACCAGAAGCTTGATGTG 1026
              || |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        689 TCGTGGAAAATTGATTCCCTGTCTTTGAACAGGGAGTTTGGGTACCAGAAGCTTGATGTG 748

Qy       1027 AAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCATCA 1086
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        749 AAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAACTCATCA 808

Qy       1087 GGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAA 1124
              ||| |||||||||||||||| ||||||||||||||
Db        809 GGACGGAAGCACAGGGGCCACTCAGAGTCGCCCGAGAA 846


RESULT 14
BD146965
LOCUS       BD146965                730 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Primer for synthesizing full-length cDNA and use thereof.
ACCESSION   BD146965
VERSION     BD146965.1  GI:27852723
KEYWORDS    JP 2002191363-A/1808.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 730)
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primer for synthesizing full-length cDNA and use thereof
  JOURNAL   Patent: JP 2002191363-A 1808 09-JUL-2002;
            HELIX RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002191363-A/1808
            PD   09-JUL-2002
            PF   28-JUL-2000 JP 2000280990
            PI   TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
            PI   SAITO,
            PI   JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
            PI   KEIICHI NAGAI,TETSUJI OTSUKI
            PC
            C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
             10,
            PC   C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
             Primer for synthesizing full-length cDNA and use thereof FH   Key
                     Location/Qualifiers
            FT   source          1. .730
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .730
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           35.9%;  Score 639.8;  DB 6;  Length 730;
  Best Local Similarity   99.0%;  Pred. No. 2e-146;
  Matches  663;  Conservative    0;  Mismatches    5;  Indels    2;  Gaps    2;

Qy          1 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 122

Qy         61 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 182

Qy        121 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 242

Qy        181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 302

Qy        241 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 362

Qy        301 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 422

Qy        361 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 482

Qy        421 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 542

Qy        481 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 602
```

```
Qy        541 GATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 600
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 GATTTGGACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 662

Qy        601 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 660
              || |||||||| || |||||||||||||||||||||||||||| || |||||||| ||||
Db        663 CANAACGTGA-CTNCTTGCTGAAGGAGTCCACGCAAGGAGTGANCA-CCTGTTCAAGGAG 720

Qy        661 ATCACAGCCT 670
              ||||||||||
Db        721 ATCACAGCCT 730


RESULT 15
AX866903
LOCUS       AX866903                730 bp    DNA     linear   PAT 17-DEC-2003
DEFINITION  Sequence 1808 from Patent EP1074617.
ACCESSION   AX866903
VERSION     AX866903.1  GI:40021252
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primers for synthesising full-length cDNA and their use
  JOURNAL   Patent: EP 1074617-A 1808 07-FEB-2001;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .730
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             35.9%;  Score 639.8;  DB 6;  Length 730;
  Best Local Similarity   99.0%;  Pred. No. 2e-146;
  Matches  663;  Conservative    0;  Mismatches    5;  Indels    2;  Gaps    2;

Qy          1 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ATGAGCGGATCACAGAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 122

Qy         61 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GTGAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGGAGATTGCCTCGGATTCCGACAGC 182

Qy        121 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 AGGGTCACCTGTCTGTGTGCCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGT 242

Qy        181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACC 302

Qy        241 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GAAACAGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCCTCAACAAGCACGAGCTGCAG 362

Qy        301 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGGTCGCGCCTGGCTGCGC 422

Qy        361 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 482

Qy        421 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 AGGCTGAGCACTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATG 542

Qy        481 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 CTTCCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTGCGATTAACATCGACAACAAG 602

Qy        541 GATTTGAACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 GATTTGGACGGGCAGAGTAAGTTTGCTCCCACCGTTTCAGACCTCTTAAAGGAGTCAACG 662

Qy        601 CAGAACGTGACCTCCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAG 660
              || |||||||| || |||||||||||||||||||||||||||| || |||||||| ||||
Db        663 CANAACGTGA-CTNCTTGCTGAAGGAGTCCACGCAAGGAGTGANCA-CCTGTTCAAGGAG 720

Qy        661 ATCACAGCCT 670
              ||||||||||
Db        721 ATCACAGCCT 730


Search completed: October 26, 2005, 04:09:03
Job time : 7729 secs
```